ISSN 1814-9316

Learning and Becoming in Practice

THE INTERNATIONAL CONFERENCE OF THE LEARNING SCIENCES (ICLS) 2014

PROCEEDINGS VOLUME 1



ICLS 2014 University of Colorado Boulder
June 23-27, 2014

www.isls.org/icls2014

HOSTED BY:

University of Colorado Boulder

SPONSORED BY:

ISLS
International Society of the Learning Sciences

Learning and Becoming in Practice:
The International Conference of the Learning Sciences (ICLS) 2014

Volume 1

11th International Conference
of the Learning Sciences

June 23-27, 2014, Colorado, USA
The University of Colorado Boulder

Editors:
Joseph L. Polman, Eleni A. Kyza, D. Kevin O'Neill, Iris Tabak,
William R. Penuel, A. Susan Jurow, Kevin O'Connor,
Tiffany Lee, and Laura D'Amico

Learning and Becoming in Practice: The International Conference of the Learning Sciences (ICLS) 2014
June 23–27, 2014
Boulder, Colorado, USA.

© 2014 International Society of the Learning Sciences, Inc. [ISLS]. Rights reserved.



www.isls.org

ISBN 978-0-9903550-1-4 (Volume 1, PDF Version)
ISSN: 1814-9316

Cite as: Polman, J. L., Kyza, E. A., O'Neill, D. K., Tabak, I., Penuel, W. R., Jurow, A. S., O'Connor, K., Lee, T., and D'Amico, L. (Eds.). (2014). *Learning and becoming in practice: The International Conference of the Learning Sciences (ICLS) 2014, Volume 1*. Boulder, CO: International Society of the Learning Sciences.

All rights reserved. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear the copyright notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior written permission of the International Society of the Learning Sciences. The International Society of the Learning Sciences is not responsible for the use which might be made of the information contained in this book.

Conference logo design by Rebecca Cober and Raymond Johnson
Cover design by Raymond Johnson

Sponsors

ICLS 2014 would like to thank our sponsors:

Cambridge University Press (cambridge.org)

Center for the Advancement of Informal Science Education (CAISE; informalscience.org)

Inquirium, LLC (www.inquirium.net)

International Society of the Learning Sciences (www.isls.org)

MacArthur Foundation (www.macfound.org)

National Center for Atmospheric Research (ncar.ucar.edu)

National Science Foundation (www.nsf.gov), sponsors of Early Career Workshop and Doctoral Consortium (Grant # DRL-1346644) and Research-Practice Partnership Workshop (Grant #DRL-1408510)

Springer Science+Business Media (www.springer.com), sponsors of Early Career Workshop

University of Colorado Boulder (www.colorado.edu)
- School of Education
- Department of Psychology and Neuroscience
- Department of Computer Science
- Center for STEM Learning
- Institute of Cognitive Science
- Provost Russell Moore

 © ISLS

CAMBRIDGE

Forthcoming Scholarship from

Cambridge University Press

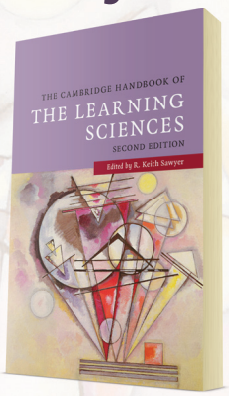
The Cambridge Handbook of the Learning Sciences, 2nd Edition

Edited by R. Keith Sawyer
University of North Carolina, Chapel Hill

Paperback (ISBN-13: 9781107626577) | $65.00 | £40.00
Hardback (ISBN-13: 9781107033252) | $160.00 | £100.00

Cambridge Handbooks in Psychology

Not yet published - available from October 2014 | 750 pages



The interdisciplinary field of the learning sciences encompasses educational psychology, cognitive science, computer science, and anthropology, among other disciplines. *The Cambridge Handbook of the Learning Sciences*, first published in 2006, is the definitive introduction to this innovative approach to teaching, learning, and educational technology. This dramatically revised second edition incorporates the latest research in the field, includes twenty new chapters on emerging areas of interest, and features contributors who reflect the increasingly international nature of the learning sciences. The authors address the best ways to design educational software, prepare effective teachers, organize classrooms, and use the internet to enhance student learning. They illustrate the importance of creating productive learning environments both inside and outside school, including after-school clubs, libraries, museums, and online learning environments. Accessible and engaging, the Handbook has proven to be an essential resource for graduate students, researchers, teachers, administrators, consultants, educational technology designers, and policy makers on a global scale.

Key Features

Written by the leading scholars in the field

Provides a comprehensive overview of the latest research

Accessible to newcomers to the field

cambridge.org/psychology @CambUP_Psych facebook.com/CambridgePsych

CAMBRIDGE UNIVERSITY PRESS

Conference Organizers and Committees

Conference Chairs
William R. Penuel - *University of Colorado Boulder, USA*
A. Susan Jurow - *University of Colorado Boulder, USA*
Kevin O'Connor - *University of Colorado Boulder, USA*

Program Chairs
Joseph L. Polman - *University of Colorado Boulder, USA*
Eleni A. Kyza - *Cyprus University of Technology, Cyprus*
D. Kevin O'Neill - *Simon Fraser University, Canada*
Iris Tabak - *Ben-Gurion University of the Negev, Israel*

Workshop Chairs
Victor Lee - *Utah State University, USA*
Yannis Dimitriadis - *University of Valladolid, Spain*

Doctoral Consortium Chairs
Barry Fishman - *University of Michigan, USA*
Mimi Recker - *Utah State University, USA*

Early Career Workshop Chairs
Naomi Miyake - *University of Tokyo, Japan*
Chris Quintana - *University of Michigan, USA*

Communications Chair
Vanessa Svihla - *University of New Mexico, USA*

Special Sessions Chairs
Kris Gutiérrez - *University of Colorado Boulder, USA*
Victoria Hand - *University of Colorado Boulder, USA*

Organizational and Administrative Support
Tiffany Lee - *University of Colorado Boulder, USA*
Joanna Weidler-Lewis - *University of Colorado Boulder, USA*
Sam Severance - *University of Colorado Boulder, USA*
Patty McDonald - *University of Colorado Boulder, USA*
Sara McDonald - *University of Colorado Boulder, USA*
University of Colorado Boulder Conference Services
Laura D'Amico – *Simon Fraser University, Canada*

Advisory Committee

Sasha Barab - *Arizona State University, USA*
Philip Bell- *University of Washington, USA*
Hilda Borko - *Stanford University, USA*
Paul Cobb - *Vanderbilt University, USA*
Kevin Crowley - *University of Pittsburgh, USA*
Danny Edelson - *National Geographic Society, USA*
Mike Eisenberg - *University of Colorado Boulder, USA*
Rogers Hall - *Vanderbilt University, USA*
Leslie Herrenkohl - *University of Washington, USA*
Ilana Horn - *Vanderbilt University, USA*
Patrick Jermann - *École Polytechnique Federale de Lausanne, France*
Yasmin Kafai - *University of Pennsylvania, USA*
Manu Kapur - *National Institute of Education, Singapore*
Paul Kirschner - *Open University of the Netherlands, Netherlands*
Moseli Mafa - *Lesotho College of Education, South Africa*
Na'ilah Nasir - *University of California at Berkeley, USA*
Miguel Nussbaum - *Pontifica Universidad Católica de Chile, Chile*
Roy Pea - *Stanford University, USA*
Jim Pellegrino - *University of Illinois Chicago, USA*
Peter Reimann - *University of Sydney, Australia*
Nikol Rummel - *Ruhr-Universität Bochum, Germany*
Anna Sfard - *University of Haifa, Israel*

Program Committee

Nancy Ares - *University of Rochester, USA*
Flávio Azevedo - *University of Texas at Austin, USA*
Angela Calabrese Barton - *Michigan State University, USA*
Katerine Bielaczyc - *Clark University, USA*
Karin Brodie - *Wits University in Johannesburg, South Africa*
Tak-Wai Chan - *National Central University of Taiwan, Taiwan ROC*
Cynthia Carter Ching - *University of California at Davis, USA*
Joshua Danish - *Indiana University, USA*
Ingrid de St. Georges - *University of Luxembourg, Luxembourg*
Noel Enyedy - *University of California at Los Angeles, USA*
Indigo Esmonde - *University of Toronto, Canada*
Deborah Fields - *Utah State University, USA*
Noah Finkelstein - *University of Colorado Boulder, USA*
Shawn Ginwright - *San Francisco State University, USA*
Melissa Gresalfi - *Vanderbilt University, USA*
Steve Guberman - *Science Museum of Minnesota, USA*
Susan Goldman - *University of Illinois at Chicago, USA*
Yael Kali - *University of Haifa, Israel*
Eric Klopfer - *Massachusetts Institute of Technology, USA*
Julie Libarkin - *The Ohio State University, USA*
April Luehmann - *University of Rochester, USA*
Leilah Lyons - *University of Illinois at Chicago, USA*
Cathy Manduca - *Carleton College, USA*
Tom Moher - *University of Illinois at Chicago, USA*
Line Mørck - *The Danish School of Education - Aarhus University (DPU), Denmark*
Eduardo Mortimer - *Universidade Federal de Minas Gerais, Brazil*
Valerie Otero - *University of Colorado Boulder, USA*
Jrene Rahm - *Université de Montréal, Canada*
Aria Razfar - *University of Illinois at Chicago, USA*
John Rogers - *University of California at Los Angeles, USA*
Beth Rubin - *Rutgers University, USA*
Ann Ryu - *University of Maryland, USA*
William Sandoval - *University of California at Los Angeles, USA*
Susan Singer - *Carleton College, USA*
Joi Spencer - *University of San Diego, USA*
Tamara Sumner - *University of Colorado Boulder, USA*
Deborah Tatar - *Virginia Tech, USA*
Carrie Tzou - *University of Washington-Bothell, USA*
Jennifer Vadeboncoeur - *University of British Columbia, Canada*
Philip Vahey - *SRI International, USA*
Jan van Aalst - *University of Hong Kong, Hong Kong*
Susan Yoon - *University of Pennsylvania, USA*

Reviewers

Anthony Aakre
Fouad Abd-El-Khalick
Dor Abrahamson
Louis Abrahamson
Samuel Abramovich
Andres Acher
Ugochi Acholonu
June Ahn
Shaaron Ainsworth
Fabio Akhras
Alicia Alonzo
Heejung An
Janice Anderson
Alessandro Antonietti
Golnaz Arastoopour
Nancy Ares
Christa Asterhan
Leslie Atkins
Katerina Avramides
Flavio Azevedo
Ari Bader-Natal
Elizabeth Bagley
Xin Bai
AnnMarie Baines
Michael Baker
Ryan Baker
Arthur Bakker
Sasha Barab
Judith Barak
Jacqueline Barnes
Lauren Barth-Cohen
Antonia Baumeister
Philip Bell
Brian Belland
Yifat Ben-David Kolikant
Bronwyn Bevan
Ruchi Bhanot
Heather Birch
Adélaïde Blavier
Erica Boling
Michael Bolz
Rebecca Boncoddo
Elizabeth Bonsignore
Angela Booker
Marcela Borge
Hilda Borko
Ivica Boticki
Cori Bower
Jonathan Boxerman
Tharrenos Bratitsis
Isabel Braun
Clare Brett
Leah Bricker
Karin Brodie
Scott Brown
Amy Bruckman
Stein Brunvand
Thomas Bussey
Donna Caccamise
Murat Perit Cakir
Angela Calabrese Barton
Elena Calderón-Canales
Kathleen Capo Crucet
Mia Čarapina
Adam Carberry
Monica Cardella
Aprea Carmela
Cris Castro-Alonso
Laura Cathcart
Collazos Cesar
Margaret Chan
Elizabeth Charles
Bodong Chen
Gaowei Chen
Mark Chen
Wenli Chen
Ying-Chih Chen
Britte Cheng
John Cherniavsky
Ellina Chernobilsky
Joshua Childs
Cynthia Carter Ching
Clark Chinn
Jeanne Chowning
Reffay Christophe
Stefanie Chye
Doug Clark
Tamara Clegg
Virginia Clinton
Hunter Close
Paul Cobb
Rebecca Cober
Barbara Colombo
Melissa Cook
Benny Cooper
Charlie Cox
Ulrike Cress
Charles Crook
Kevin Crowley
Barney Dalgarno
Crina Damsa
Brian Danielak
Joshua Danish
Pryce Davis
Ton De Jong
Ingrid de Saint-Georges
Cindy De Smet
Reuma De-Groot
Angela DeBarger
Chris Dede
Cesar Delgado
David DeLiema
Benjamin DeVane
Nicoletta Di Blas
Michael Dianovsky
Ning Ding
Betsy DiSalvo
Neven Drljević
Izabel Duarte Olson
Gabriella Ducamp
Therese Dugan
Sean Duncan
Shree Durga
Elizabeth B. Dyer
Gregory Dyke
Catherine Eberbach
Julia Eberle
Daniel Edelson
Ann Edwards
Michael Eisenberg
Hebbah El-Moslimany
Jan Elen
Judith Enriquez
Noel Enyedy
Gijsbert Erkens
Bernhard Ertl
Indigo Esmonde
Wanda Eugene
Michael Evans
Howard Everson
Cameron Fadjo
Kirschner Femke
Leila Ferguson
Georgios Fessakis
Jill Fielding-Wells
Deborah Fields
Noah Finkelstein
Kara Finnigan
Frank Fischer
Cory Forbes
Ellice Forman
Aroutis Foster
Ina Fourie
Erin Marie Furtak
Judi Fusco
Krista Galyen
Brian Gane
Iolanda Garcia
Alejandra Garcia-Franco
Maria Teresa Gastardo
Xun Ge
Andreas Gegenfurtner
Andrew Gibbons
Aristotelis Gkiolmas
Alessandro Gnoli
Janice Gobert
Susan R. Goldman
Kimberley Gomez
Monica Gordon Pershey
Amelia Gotwals
Julia Gouvea

 © ISLS

Jamillah Grant
Jeffrey Greene
Melissa Gresalfi
Gabriela Groza
Brigitte Gruson
Xiaoqing Gu
Steven Guberman
Bekir Gur
Mark Guzdial
Seungyon Ha
Alan Hackbarth
Iassen Halatchliyski
Rogers Hall
Tony Hall
Erica Halverson
Richard Halverson
Victoria Hand
Nazir Hawi
Kenneth Hay
Elisabeth Hayes
Katie Headrick Taylor
Andrew Heckler
Neil Heffernan
Phillip Herman
Paula Heron
Leslie Herrenkohl
Arnon Hershkovitz
Daniel Hickey
Yotam Hod
Hayat Hokayem
Nathan Holbert
Margaret Holzer
Huang-Yao Hong
Sameer Honwad
Paula Hooper
Matthew Hora
Ilana Horn
Michael Horn
Anesa Hosein
Pei-Hsuan Hsieh
Roland Hubscher
Sarah Hug
Chia Yuan Hung
Liisa Ilomäki
Birgit Imhof
Åke Ingerman
Adam Ingram-Goble
Ioan Gelu Ionas
Rebecca Itow
Sanna Järvelä
Kara Jackson
Michael Jacobson
Azilawati Jamaludin
Ellen Jameson
Allan Jeong
Heisawn Jeong
Patrick Jermann
Osvaldo Jimenez
Rick Jochen
Emily Johnson
Mina Johnson
Aditya Johri
Zywica Jolene
Michelle Jordan
Yih-Ruey Juang
A. Susan Jurow
Seokmin Kang
Gellof Kanselaar
Shulamit Kapon
Manu Kapur
Petter Karlström
Danielle Keifert
Lisa Kenyon
Khusro Kidwai
Josephine Kilde
Beaumie Kim
ChanMin Kim
Dongsik Kim
Kibum Kim
Mi Song Kim
Joachim Kimmerle
Elizabeth King
Jennifer King Chen
Paul Kirschner
Ben Kirshner
Elaine Klein
Jayne Klenner-Moore
Laure Kloetzer
Eric Klopfer
Simon Knight
Mon-Lin Ko
Marta Kobiela
Alaric Kohler
Ingo Kollar
Birgitta Kopp
Timothy Koschmann
Donna Kotsopoulos
Michail D. Kozlov
Joseph Krajcik
Moshe Krakowski
Karsten Krauskopf
Andrew Krumm
Andy Krumm
Beth Kubitskey
Kristiina Kumpulainen
Eric Kuo
Samuel Kwon
Eleni A. Kyza
Barbara Ladewski
Therese Laferriere
Minna Lakkala
Diane Lam
Jennifer Langer-Osuna
Mike Lawson
Ard Lazonder
Heather Leary
Heather Leary
Chwee Beng Lee
Jennifer Lee
KoSze Lee
Kyungmee Lee
Tiffany Lee
Victor Lee
Christopher Leeder
Rich Lehrer
Olivia Levrini
Dalit Levy
Armanda Lewis
Colleen Lewis
Katherine Lewis
Xiao Li
Chang-Yen Liao
Emily Lin
Hsien-Ta Lin
Jun Lin
Robb Lindgren
Oskar Lindwall
Jennifer Lineback
Matthew Lira
Shiyu Liu
Stanley Lo
Steven Lonn
Chee Kit Looi
Suzanna Loper
Maritza Lozano
Margaret Lucero
Sten Ludvigsen
Michelle Lui
Melissa Luna
Kristine Lund
Kristine Lund
Patrik Lundh
Guanzhong MA
Jasmine Ma
Augusto Macalalag Jr.
Kati Maekitalo-Siegl
Moseli Mafa
Alejandra Magana
Alecia Marie Magnifico
Alan Maloney
Nils Malzahn
Stefania Manca
Nicole Mancevice
Andrew Manches
Michael Manderino
Cathyrn Manduca
Eve Manz
Bairral Marcelo
Specht Marcus
Thomas Martens
Crystle Martin
Fred Martin
John Martin
Wessner Martin
Roberto Martinez Maldonado
Patricia Marzin
Anastasios Matos
Camillia Matuk
Lucy McClain
Scott McDonald

 © ISLS

Steven McGee
Veronica Cassone McGowan
Susan McKenney
Katherine McNeill
Jenna McWilliams
Richard Medina
José Melendez
Emma Mercier
Joi Merritt
Joi Merritt
Georgia Michael
Jacobson Michael
Vera Michalchik
Toshio Mochizuki
Tom Moher
Emily Moore
Anders Morch
Line lerche Morck
Judit Moschkovich
Johannes Moskaliuk
Chrystalla Mouza
Jin Mu
Hanni Muukkonen
Khine Myint Swe
Matthias Nückles
Denise Nacu
Na'ilah Nasir
John Nesbit
Kenny Nguyen
Lijun Ni
Iolie Nicolaidou
Kraemer Nicole
Nicolae Nistor
Omid Noroozi
Michele Notari
Iskra Nunez
Adaeze Nwaigwe
Kevin O'Connor
Angela O'Donnell
Kevin O'Neill
Amanda Ochsner
Sandra Okita
Chandra Orrill
Jun Oshima
Valerie Otero
Hiroki Oura
Sharon Oviatt
Saliha Ozkan-Bekiroglu
Annemarie Palincsar
Chrystalla Papademetri-Kachrimani
Orit Parnafes
Frederick Peck
Margus Pedaste
William Penuel
Tana Peterman
Erin Pfaff
Rachel Phillips
Anthony Phonethibsavads
Vicky Pilitsis
Julia Plummer
William Pluta
Joe Polman
David Porcaro
Torsten Porsch
Aaron Price
Naomi Prusak
Burke Quinn
Ritu Radhakrishnan
Jrene Rahm
Antti Rajala
Aria Razfar
Janet Read
Mimi Recker
Stephen Reed
Richard Reeve
Martin Rehm
Peter Reimann
K. Ann Renninger
Suzanne Rhodes
Rocco Ricci
Gabriela Richard
Christoph Richter
Antti Rissanen
Giuseppe Ritella
Kelsey Rodgers
Ido Roll
Margarida Romero
Jeremy Roschelle
Carolyn Rosé
Elizabeth Rowe
Beth Rubin
Nikol Rummel
Donna Russell
Stephanie Ryan
Minjung Ryu
Nora Sabelli
Ornit Sagy
Ron Salden
Asmalina Saleh
William Sandoval
Rafi Santo
Elissa Sato
Vashti Sawtelle
Keith Sawyer
Eleanor Sayre
Giovanna Scalone
Andrea Scheer
Rachel Scherr
Annett Schmeck
Bertrand Schneider
Emily Schoerning
Kathy Schuh
Stephan Schwan
Neil Schwartz
Baruch Schwarz
Jennifer Schwarz
Beat Schwendimann
Stephanie Scopelitis
Kimberly Scott
Jen Scott Curwood
James Segedy
Colleen Seifert
Anna Sfard
Li Sha
David Shaenfield
Niral Shah
Tara Rosenberger Shankar
R. Benjamin Shapiro
Priya Sharma
Erin Shaw
Nicole Shea
Brett Shelton
Ji Shen
Bruce Sherin
Mayumi Shinohara
Wesley Shumar
Bernadette Sibuma
Tiffany-Rose Sikorski
Tiffanyrose Sikorski
Jean Simon
Tuya Siqin
Irene Skuballa
Brian Slattery
Bert Slof
Jim Slotta
Ole Smørdal
Brian Smith
Brian Smith
Debbie Smith
Erica Snow
Minyoung Song
Firat Soylu
Hans Spada
Sarah Stager
Gerry Stahl
Trausan-Matu Stefan
Karsten Stegmann
Daniel Steinbock
Jan-Willem Strijbos
Shelley Stromholt
Andrew Stull
Florence Sullivan
Tamara Sumner
Shannon Sung
Patrick Sunnen
Daniel Suthers
Vanessa Svihla
Karen Swan
Richard Swan
Antti Syvänen
Iris Tabak
Lori Takeuchi
Miwa Takeuchi
Kok-Sing Tang
Pierre Tchounikine
Chakorn Techatassanasoontorn
Meng Yew Tee
Takao Terano
Pattaraporn Thampradit

 © ISLS

Alexander Thayer
Nicolet Theunissen
Jakob Tholander
Mike Tissenbaum
Sachiko Tosa
Michael Tscholl
Dimitra Tsovaltzi
Hakan Tuzun
Kersti Tyson
Dora Tzialli
Carrie Tzou
Nan Uhlik
Jessica Umphress
Josh Underwood
Jennifer Vadeboncoeur
Phil Vahey
Jan van Aalst
Marije van Amelsvoort
Carla van de Sande
Beth van Es
Katie Van Horne
Wouter van Joolingen
Herre van Oostendorp
Charles Vanover
Keisha Varma
Laura Varner
Richard Vath
Marjo Virnes
Joke Voogt
Chieu Vu Minh
Hanna Vuojärvi
Aditi Wagh
Richard Walker
Candace Walkington
Robert Wallon
Jeremy Wang
X. Christine Wang
Peter Wardrip
Muffie Waterman
Dana Webber
Anne Weibert
Jennifer Weible
Joanna Weidler-Lewis
Armin Weinberger
Michael Weinstock
Yun Wen
Daniel Wessel
Astrid Wichmann
Caroline Williams
Robert Williams
Margaret J. Wilsman
Kerri Wingert
Phil Winne
Alyssa Wise
Lung-Hsiang Wong
Iwan Wopereis
Kenneth Wright
Kenneth Wright
Hsin-Kai Wu
Longkai Wu
Qiong Wu
Ruth Wylie
Sihan Xiao
Ying Xie
Kimberly Xu
Kali Yael
Senom Yalcin
Anat Yarden
Yau-yuen Yeung
Jason Yip
Susan Yoon
Yawen Yu
Carmen Zahn
Raul Zaritsky
Lin Zhang
Yonghe Zhang
Zhihui Helen Zhang
Heather Toomey Zimmerman
Bill Zoellick
Steve Zuiker

 © ISLS

Preface

The international and interdisciplinary field of the Learning Sciences brings together researchers from the fields of cognitive science, educational research, psychology, computer science, artificial intelligence, information sciences, anthropology, sociology, neurosciences, and other fields to study learning in a wide variety of formal and informal contexts (see www.isls.org). This field emerged in the late 1980s and early 1990s, with the first International Conference of the Learning Sciences (ICLS) held in 1991 at Northwestern University in Evanston, Illinois, USA. Subsequent meetings of ICLS were held again in Evanston, USA (1996), Atlanta, GA, USA (1998), Ann Arbor, MI, USA (2000), Seattle, WA, USA (2002), Santa Monica, CA, USA (2004), Bloomington, IN, USA (2006), Utrecht, the Netherlands (2008), Chicago, IL, USA (2010), and Sydney, NSW, Australia (2012). The 11th ICLS meeting in 2014 was hosted at the University of Colorado Boulder, USA.

Papers for this conference were submitted in November 2013, and then went through a process of peer review. We received a record number of submissions (749), 50% more than for any past ICLS conference. Overall, 306 submissions were accepted, which is an 18% increase from previous conferences. The overall acceptance rate for submissions was 41%.

Acceptance rates for each category were:

- 32% for full papers
- 38% for reports/reflections
- 52% for posters
- 55% for symposia

The program reflects broad geographic representation, with contributions from 21 countries on 5 continents.

We are especially grateful to those who performed reviews. A total of 610 people completed over 2,300 reviews of the submissions. As in recent years, for each symposium and full paper, we assigned a senior reviewer who examined all reviews and made a recommendation regarding acceptance in the category submitted, acceptance in another category, or rejection. These senior reviewers greatly helped us make decisions on acceptance for each submission.

The theme of ICLS 2014 is "Learning and Becoming in Practice." By focusing on learning and *becoming*, we aimed to foreground the ways that learning entails becoming a certain kind of person. By focusing on learning and becoming *in practice*, we aimed to foreground the ways that learning processes are situated within different kinds of practices. We identified three kinds of practices that capture the range of contexts and processes in which people can learn: by engaging in the epistemic practices of disciplines, by participating in sociocultural practices, and by engaging in design. Two additional practices we highlight pertain to how we organize our own work as learning scientists: the practices for analyzing and modeling learning across settings and time, and the practices for designing for scale and sustainability.

In many ways, practice is a natural focus for our field. The call for conducting design research grew in part from a perception that findings generated in laboratory studies of cognition answered only a subset of the questions we had about learning. Design researchers take a deeply pragmatic stance toward research on learning, seeking to generate insights from studying learning in specific contexts. People who collaborated within key institutions in the

history of the learning sciences—such as the Institute for Research on Learning and Xerox PARC—were key to developing the rich and generative theoretical accounts of learning in practice.

The different strands related to the theme of "learning and becoming in practice" highlight several lines of inquiry in the learning sciences that address five key questions, which we elaborate below.

How Do People Learn Core Disciplinary Ideas by Engaging in Epistemic Practices of Disciplines?

By disciplines, we include not only the learning in K-12 school science and mathematics, which makes up the majority of learning sciences research, but also learning in higher education and in other disciplines, including engineering, social sciences, and the humanities. The term epistemic practices refers to how different disciplines argue that people come to know and warrant their ideas; the study of learning in epistemic practices encompasses how people come to be able to participate in these practices. Scholars often speak of the epistemic "commitments" that define the boundaries of particular disciplinary communities, and this idea of commitments signals how people must come to understand and appropriate particular norms for thinking, speaking, and reasoning to be part of that community. Contemporary learning sciences research on epistemic practices is wide-ranging and includes studies of how children's understandings of the practices of modeling in science develop over time, as well as studies of classroom discourse practices and teachers' orchestration of them. Research has also highlighted how young people navigate between everyday and disciplinary forms of knowing in ways that shape their identities. Learning sciences research has also explored how such epistemic practices as explanation develop within family conversations and museums, as well as how the everyday epistemologies of learners from nondominant groups relate to epistemic practices of the disciplines.

How Do People Learn through Participation in Sociocultural Practices?

The landmark volume, *How People Learn*, synthesized decades of research on learning and has greatly informed how educators design learning environments. Several of the committee members who were involved in that effort have since called for a second volume, focused on the idea of "how people learn culturally." In emphasizing culture, they draw attention to something that the report included but was not in the foreground, namely that learning is a deeply social and cultural process. Studies of learning within sociocultural practices often draw on Vygotskian and neo-Vygotskian theories of learning and development, but not exclusively so. Studies of cultural cognition in psychology and anthropology have made and continue to provide important insights into learning, as do experimental studies of social and cultural aspects of learning. Our purpose in posing this question as a central strand for our conference theme was to encourage dialogue and attention to this methodologically and theoretically diverse body of work in the field.

How and What Do People Learn by Engaging in Practices of Design?

Our field has a rich tradition of research, especially within science and engineering studies, of design as a way to learn. By participating in design, learners engage deeply with disciplinary and related content; when they do it with others, they also gain practice in the valued skills of collaboration and teamwork. In the past decade, within and along the periphery of the learning sciences, the scope of what learners design has expanded. Many projects are

exploring what youth learn, for example, when they engage in complex activities of media production or contribute to social media. Still others are engaged in innovative 3D and technology-based physical construction.

An ongoing conversation within the field focuses on design-based research as a methodology. By no means settled is the debate over what it means to warrant claims about what we learn from engaging in this form of research. Other methodologies, too, play a central role in our field—from critical ethnographies to in-depth analyses of classroom discourse—that do not involve design per se. Yet these same methodologies also have promise to help us understand more about what we learn from engaging in design and how we come to know it, especially when applied to the study of our practices of design-based research.

What Practices Should We Use to Model and Analyze Learning across Time and Across Settings?

A number of us in the community are engaged in innovative efforts to model and analyze learning over time and across settings. Our foci and approaches vary widely. We have conducted micro-analyses of learning using fine-grained knowledge analysis approaches, conducted longer-term developmental analyses, and mapped learning progressions within disciplines. Some in the fields of data mining and learning analytics are engaged in efforts to construct models from large data sets of learning pathways through specific content, especially in online learning environments. Still others are engaged in ethnographic studies of learning across settings and time. Some of these researchers are specifically focused on the roles of space and place within learning. Investigators across these different lines of research employ very different assumptions about the nature of learning, which makes the opportunity ripe for dialogue about the assumptions underlying the different approaches.

How Can We Transform Our Practice to Design More Effectively for Scaling and Sustainability?

Many learning scientists aim to have broad impact on the fields of practice that we study, whether those are schools, museums, or another setting for learning. At the same time, we recognize that limited funding and poor infrastructure hamper our efforts to achieve such an impact. We also know that by selecting environments that are more "felicitous" for design, such as well-resourced school districts, we can unwittingly exacerbate problems associated with equity of access to powerful learning opportunities. Hence, we want to engage the community in a dialogue about how we might design more effectively for scaling and sustainability, which will provide an opportunity to highlight a wide range of efforts within the field, from rapid prototyping of online environments to emerging efforts to undertake design research at the district level, not just within classrooms or individual out-of-school settings. A key theme of many of the contributions to the conference is the importance of engaging practitioners at different levels of educational systems in design, as a means to promote more transformational and sustainable changes within systems. This dialogue allows us to pursue questions, too, of how we might need to engage in efforts to re-organize systems of learning to promote more equitable learning outcomes for all.

In these proceedings volumes, you will find a wide variety of approaches to the above questions, and we look forward to continuing the conversations these papers and sessions initiated at the conference.

Finally, we express our deepest gratitude to the many people who made the conference possible: the organizing committee, advisory committee, program committee, reviewers, sponsors, volunteers, staff, and all conference presenters and participants. Your contributions make the learning sciences a thriving field, striving to transform learning opportunities that enable people to become agents of change in their own and others' lives.

- Program Chairs Joseph L. Polman, Eleni A. Kyza, D. Kevin O'Neill, and Iris Tabak
  Conference Chairs William R. Penuel, A. Susan Jurow, and Kevin O'Connor

 © ISLS

Table of Contents

Volume 1

Keynotes

Changing Practice 3
Jean Lave, Rogers Hall

Research-Practice Partnerships 4
Megan Bang, Paul Cobb, Kara Jackson, Michael Sorum, Kris D. Gutiérrez

Designing with Communities: Transforming Historically Powered Relations in Teaching and Learning 4
Megan Bang

Partnering with School and Districts to Support All Students' Learning 4
Paul Cobb, Kara Jackson, Michael Sorum

I, Thou, and Them: Distributed Memory and Learning 6
Geoffrey C. Bowker, Kevin O'Connor

Approaches to Studying and Modeling Learning Across Setting and Time 7
Jeremy Roschelle, Anna Sfard, Reed Stevens, Leona Schauble, Beth Warren, Marianne Wiser

Invited Sessions

Presidential Session: Learning and Assessment of 21st Century Skills 11
Eleni A. Kyza, Jan van Aalst, Daniel T. Hickey, Alina von Davier, Jan-Willem Strijbos, Cindy E. Hmelo-Silver

Research-Practice Partnerships in Communities 12
Daniel J. Gallagher, Nichole Pinkard, Jasmine Alfonso-Gurneau, Megan Bang, Mary Dempsey, Lori Faber, Ananda Marin, Tim S. Truitt

Children Becoming Collaborators 13
Barbara Rogoff, Rebeca Mejía-Arauz, Lucía Alcalá, Andrew Coppens, Andy Dayton, Angélica López, Omar Ruvalcaba, Yolanda Corona Caraveo, Maricela Correa-Chávez, Kris D. Gutiérrez, Luis Urrieta

Teachers as Designers 14
Joke Voogt, Susan McKenney, Yael Kali, Alain Breuleux, Rebecca Cober, James D. Slotta, Bat-Sheva Eylon, Rebecca Itow, Karen Könings, Therese Laferrière, Marcia C. Linn, Lina Markauskaite, Richard Reeve, Ornit Sagy, Hyo-Jeong So, Vanessa Svihla, Esther Tan, Camillia Matuk

Where are the Learning Sciences in the MOOC Debate? 15
George Siemens, Pierre Dillenbourg, Gerhard Fischer, Danielle S. McNamara, Nikol Rummel

Learning and Becoming through Making and Participatory Media 18
Kristiina Kumpulainen, Julian Sefton-Green, Karen Brennan, Anna Mikkola, Kylie Peppler, Elisabeth Soep

Papers

Leveling Transparency via Situated Intermediary Learning Objectives (SILOs) 23
Dor Abrahamson, Kiera Chase, Vishesh Kumar, Rishika Jain

Supporting Middle Schoolers' Use of Inquiry Strategies For Discovering Multivariate Relations In Interactive Physics Simulations 31
Luke D. Conlin, Nicole R. Hallinen, Daniel L. Schwartz

 © ISLS

Klauer's Inductive Reasoning Training as a Cognitive Apprenticeship Approach for Special-Needs Students 38
Antonia E. E. Baumeister, Heiner Rindermann

Development of an Empirically-Based Learning Performances Framework for 3rd-Grade Students' Model-Based Explanations about Hydrologic Cycling 46
Cory T. Forbes, Christina V. Schwarz, Laura Zangori

Problematizing as Scaffold for Engaging in Scientific Argumentation 54
Mon-Lin Monica Ko

Three Diagnoses of Why Transfer Across Disciplines Can Fail and Their Implications for Interdisciplinary Education 62
Eric Kuo, Danielle Champney

Diving into Practice with Children and Undergraduates: A Cultural Historical Approach to Instatiating Making and Tinkering Activity in a Designed Learning Ecology 70
Lisa Hope Schwartz, Daniela K. DiGiacomo, Kris D. Gutiérrez

Facilitating Design Research by Mapping Design Research Trajectories 78
Guanzhong Ma, Jan van Aalst

Connected Gaming: Towards Integrating Instructionist and Constructionist Approaches in K-12 Gaming 86
Yasmin B. Kafai, Quinn Burke

Learning With Multiple Visualizations in the Science Museum 94
Joyce Wang, Susan Yoon

The Role of Inconsistencies in Collaborative Knowledge Construction 102
Martina Bientzle, Ulrike Cress, Joachim Kimmerle

More Than Just Plain Old Technology Adoption: Understanding Variations in Teachers' Use of an Online Planning Tool 110
Heather Leary, Victor R. Lee, Mimi Recker

Using an Adaptive Expertise Lens to Understand the Quality of Teachers' Classroom Implementation of Computer-Supported Reform Curricula in High School Science 118
Susan Yoon, Jessica Koehler, Joyce Wang, Emma Anderson, Eric Klopfer

Promoting 5th Graders' Views of Science and Scientific Inquiry in an Epistemic-Enriched Knowledge-Building Environment 126
Feng Lin, Carol K. K. Chan, Jan van Aalst

Becoming an Activist-Mathematician in an Age of Austerity 134
Indigo Esmonde, Joe Curnow, Dominique Riviere

Measuring Affective Experience in the Midst of STEM Learning 142
Jayson Nissen, Jonathan Shemwell

Epistemic Networks for Epistemic Commitments 150
Simon Knight, Golnaz Arastoopour, David Williamson Shaffer, Simon Buckingham Shum, Karen Littleton

Framing Reflections on Instruction: A Precursor to Noticing 158
Vicky Pilitsis, Ravit Golan Duncan

"Teach me how to Facebook!" Design-Based Research about Risk Prevention on Social Network Sites 166
Ellen Vanderhoven, Tammy Schellens, Martin Valcke

The Role of Identity Development within Tensions in Ownership of Science Learning 174
Jason Yip, Tamara Clegg, June Ahn, Elizabeth Bonsignore, Michael Gubbels, Emily Rhodes, Becky Lewittes

Developing Mechanistic Model-Based Explanations of Phenomena: Case Studies of Two Fifth Grade Students' Epistemologies in Practice Over Time 182
Christina V. Schwarz, Li Ke, May Lee, Joshua Rosenberg

MOOCs: A Perspective from the Learning Sciences 190
Michael Eisenberg, Gerhard Fischer

How Interpreters Make Use of Technological Supports in an Interactive Zoo Exhibit 198
Brian Slattery, Leilah Lyons, Priscilla Jimenez Pazmino, Brenda A. López Silva, Tom Moher

Becoming an Activist: Learning the Politics and Performances of Youth Activism Through Legitimate Peripheral Participation 206
Joe Curnow

"We should all help each other": Latina undergraduates' practices and identities in the figured world of computing 214
Heather Thiry, Sarah Hug

Identifying Transfer of Inquiry Skills across Physical Science Simulations using Educational Data Mining 222
Michael Sao Pedro, Yang Jiang, Luc Paquette, Ryan S. Baker, Janice Gobert

Reactivation of Multimodal Representations and Perceptual Simulations for Meaningful Learning: A Comparison of Direct Embodiment, Surrogate Embodiment, and Imagined Embodiment 230
Saadia A. Khan, John B. Black

Creativity as Practice(d) in a Design Studio 238
Christoph Richter, Julia Lembke, Elisa Ruhl, Heidrun Allert

Fostering Scientific Reasoning. A Meta-analysis on Intervention Studies 246
Katharina Engelmann, Frank Fischer

Knowledge Organization with Multiple External Representations for Socioscientific Argumentation: A Case on Nuclear Energy 254
Bahadir Namdar, Ji Shen

Students' Resources for the Construction of Scales for Graphing 262
Cesar Delgado, Margaret Lucero

Scientific Practices Through Students' Eyes: How Sixth Grade Students Enact and Describe Purposes for Scientific Modeling Activities Over Time 270
Christina Krist, Brian J. Reiser

The Impact of a Social Robot's Attributions for Success and Failure in a Teachable Agent Framework 278
Kasia Muldner, Victor Girotto, Cecil Lozano, Winslow Burleson, Erin A. Walker

Characterizing a New Dimension of Change in Attending and Responding to the Substance of Student Thinking 286
Jennifer Richards, Andrew Elby, Ayush Gupta

Spatial Practices in CSCL Discussions 294
Benzi Slakmon, Baruch B. Schwarz

Collective Immersive Simulations: A New Approach to Learning and Instruction of Complex Biology Topics 301
Michelle Lui, James D. Slotta

A Study of Subjective Emotions, Self-Regulatory Processes, and Learning Gains: Are Pedagogical Agents Effective in Fostering Learning? 309
Nicholas Mudrick, Roger Azevedo, Michelle Taub, Reza Feyzi Begnagh, François Bouchet

Design-Based Research Process: Problems, Phases, and Applications 317
Matthew W. Easterday, Daniel Rees Lewis, Elizabeth M. Gerber

Modeling the Construction of Evidence through Prior Knowledge and Observations from the Real World 325
Lauren Barth-Cohen, Daniel Capps, Jonathan Shemwell

Automatic Coding of Questioning Patterns in Knowledge Building Discourse 333
Jin Mu, Jan van Aalst, Carol K. K. Chan, Ella Fu

Enacted Misconceptions: Using Embodied Interactive Simulations to Examine Emerging Understandings of Science Concepts 341
Robb Lindgren, Michael Tscholl

Friendship, Participation, and Site Design in Interest-Driven Learning among Adolescents 348
Ashley Cartun, Ben Kirshner, Emily Price, Adam York

Using Models for Reasoning and Content Learning: Patterns of Bootstrapping Towards Earth Science Understandings 354
Ann E. Rivet, Cheryl A. Lyons, Alison R. Miller

Towards a Complex Systems Meta-Theory of Learning as an Emergent Phenomenon: Beyond the Cognitive Versus Situative Debate 362
Michael J. Jacobson, Manu Kapur, Peter Reimann

Adventures in Argument: Training in Argumentation Influences Student Resource Use in Collaborative Meaning Making 370
Julia Gressick, Sharon J. Derry

Collective Engagement in a Technologically Mediated Science Learning Experience: A Case study in a botanical garden 378
Fariha H. Salman, Heather Toomey Zimmerman, Susan M. Land

Promoting Student Learning through Automated Formative Guidance on Chemistry Drawings 386
Anna N. Rafferty, Libby Gerard, Kevin McElhaney, Marcia C. Linn

Computer-Enhanced Dialogic Reflective Discourse 394
Shiri Mor-Hagani, Dani Ben-Zvi

A Case Study Examining the Microdynamics of Social Positioning within the Context of Collaborative Group Work 402
Lesley Dookie

Showing What They Know: Multimedia Artifacts to Assess learner understanding 410
Cindy E. Hmelo-Silver, Carolyn A. Maher, Marjory F. Palius, Robert Sigley, Alice Alston

Varied Appropriation of Tools from Professional Development: Moving Beyond Levels 418
Huy Q. Chung, Elizabeth A. van Es

Learning in Low-Performing School Districts: Conceptual and Methodological Challenges Resulting from Network Churn 426
Kara S. Finnigan, Alan J. Daly

Shifts in Identification in a Hybrid Space 434
Kok-Sing Tang

Shared Epistemic Agency and Agency of Individuals, Collaborative Groups, and Research Communities 440
Crina Damşa

Interrogating the Divide: A Case Study of Student Technology Use in a One-to-one Laptop School 448
Nicholas Wilson

Capturing Personal and Social Science: Technology for Integrating the Building Blocks of Disposition 455
Tamara Clegg, Elizabeth Bonsignore, June Ahn, Jason Yip, Daniel Pauw, Michael Gubbels, Becky Lewittes, Emily Rhodes

Moving Beyond Case Studies: Using Social Network Analysis to Study Learning as Participation in Communities of Practice 463
Julia Eberle, Karsten Stegmann, Frank Fischer

Designing Critique to Improve Conceptual Understanding 471
Elissa Sato, Marcia C. Linn

Being Mathematical Relations: Dynamic Gestures Support Mathematical Reasoning 479
Candace Walkington, Rebecca Boncoddo, Caroline Williams, Mitchell J. Nathan, Martha W. Alibali, Erica Simon, Elizabeth Pier

"Learning to Live": Expansive Learning and Mo(ve)ments Beyond 'Gang Exit' 487
Line Lerche Mørck

Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science 495
Niral Shah, Colleen Lewis, Roxane Caires

Hands-on Small Group versus Whole Class Use of an Interactive Simulation: Qualitative Comparisons 503
A. Lynn Stephens, John J. Clement

The Contours and Possibilities of Desire in Sociocultural Research on Learning and Becoming 511
Ian Parker Renga

How Do Learners Process Information in Lectures? The Role of Projected Slides and Type of Note-taking 519
Christof Wecker

Using Analytics for Improving Implementation Fidelity in an Large Scale Efficacy Trial 527
Mingyu Feng, Jeremy Roschelle, Robert Murphy, Neil T. Heffernan

Where Are We Now? Research Trends in the Learning Sciences 535
Elizabeth Koh, Young Hoan Cho, Imelda Caleon, Yu Wei

Defining Success in an Alternative High School: Resources for the Reframing of Education 543
Gavin Tierney

“Case n' Point”: Discovering Learning in the Nonce 551
Timothy Koschmann, Alan Zemel, Michael Neumeister

Time Needed: Growth of Preservice Science Teachers’ Knowledge of Inquiry and Practice of Lesson Design 559
Augusto Z. Macalalag Jr.

Becoming Agents of Change through Participation in a Teacher-Driven Professional Research Community 567
Michael Ross, Ben Van Dusen, Valerie Otero

‘Mangling’ Science Instruction: Creating Resistances to Support the Development of Practices and Content Knowledge 575
Eve Manz

An Analytic Tool for Supporting Teachers' Reflection on Classroom Talk 583
Gaowei Chen, Sherice N. Clarke, Lauren B. Resnick

The Discourse of Creative Problem Solving in Childhood Engineering Education 591
Elise Deitrick, Brian O'Connell, R. Benjamin Shapiro

Supporting Pre-Service Science Teachers' Planning of Task-Based Classroom Discussions 599
Danielle Ross, Aaron M. Kessler, Jennifer Cartier

Understanding the Relationships Within and Between Constructs of a Learning Progression: Combining Multidimensional Item Response Modeling and Latent Class Analysis 607
Jinnie Choi, Ravit Golan Duncan

Students' Use of Evidence and Epistemic Criteria in Model Generation and Model Evaluation 615
Ravit Golan Duncan, Carol Tate, Clark A. Chinn

Volume 2

Papers (continued)

The Role of Scientific and Social Academic Norms in Student Negotiations while Building Astronomy Models 625
Melissa Sunshine Cook, Noel Enyedy

The Beginnings of Engineering Design in an Integrated Engineering and Literacy Task 633
Mary McCormick, David Hammer

The Roles of Teacher Questioning in Argument-based Inquiry (ABI): Approaches that Promote Cognitive Thinking and Dialogical Interaction 641
Ying-Chih Chen, Brian Hand

Hear What They Say and Watch What They Do: Predicting Valid Mathematical Proofs Using Speech and Gesture 649
Elizabeth Pier, Candace Walkington, Caroline Williams, Rebecca Boncoddo, Jessica Waala, Martha W. Alibali, Mitchell J. Nathan

I Want to be a Game Designer or Scientist: Connected Learning and Developing Identities with Urban, African-American Youth 657
June Ahn, Mega Subramaniam, Elizabeth Bonsignore, Anthony Pellicone, Amanda Waugh, Jason Yip

Reflecting on Educational Game Design Principles via Empirical Methods 665
Osvaldo Jimenez

Exploring Group-Level Epistemic Cognitions within a Knowledge Community and Inquiry Curriculum for Secondary Science 673
Alisa Acosta, Michelle Lui, James D. Slotta

High School Students' Parameter Space Navigation and Reasoning during Simulation-Based Experimentation 681
Hee-Sun Lee, Amy Pallant, Robert Tinker, Paul Horwitz

Learner Alignment with Expansive Framing as a Driver of Transfer 689
Diane P. Lam, Adam Mendelson, Xenia S. Meyer, Lloyd Goldwasser

Game-Enabled Agency: Outcomes that Matter 697
Sasha Barab, Anna Arici

Model-Based Reasoning: A Framework for Coordinating Authentic Scientific Practice with Science Learning 705
Julia S. Gouvea, Arash Jamshidi, Cynthia Passmore

The Impact of Text Genre on Science Interest in an Authentic Science Learning Environment 713
Steven McGee, Amanda M. Durik, Dena Ann Pastor

When Experts Disagree: Sourcing Practices While Reading Conflicting Online Information Sources 721
Sarit Barzilai, Eynav Tzadok, Yoram Eshet-Alaklai

Learning and Becoming in an After School Program: The Relationship as a Tool for Equity within the Practices of Making and Tinkering 729
Daniela K. DiGiacomo, Kris D. Gutiérrez

"It's Intentional": Co-Construction of Transformational Processes and Pathways within and across Hubs of Interdependence in an Urban Community 737
Joanne Larson, Courtney Hanny, Joyce Duckles, Joel Gallegos Greenwich, Eric Meyer, Robert Moses, George Moses, Kimberly Jones, Jeremy Smith

Middle School Learners' Ontological 'Trying-on' of Dimensions: A Phenomenological Investigation 745
Keri Duncan Valentine, Theodore J. Kopcha

Towards the Facilitation of an Online Community of Learners: Assessing the Quality of Interactions in Yammer 753
Marcela Borge, Sean Goggins

Communities of Learning Practice: Balancing Emergence and Design in Educational Settings 761
Filitsa Dingyloudi, Jan-Willem Strijbos

Dynamic Visualization of Motion for Student-Generated Graphs 769
Jonathan M. Vitale, Kevin Lai, Marcia C. Linn

Supporting Teacher Learning for Pedagogical Innovation Through Collaborative Co-Design: Issues and Challenges 777
Nancy Law, Johnny Yuen, Yeung Lee

Inexplicable Silence: An Uncomfortable Analysis of the Social Silences 785
Daniel Steinbock

Designing for Democracy in Education: Participatory Design and the Learning Sciences 793
Betsy DiSalvo, Carl DiSalvo

Leveraging the Cultural Practices of Science for Making Classroom Discourse Accessible to Emerging Bilingual Students 800
Enrique Suárez, Valerie Otero

Using Contrasting Video Cases of the Enactment of Cognitively Demanding Science Tasks in Professional Development 808
Miray Tekkumru Kisa, Mary Kay Stein

Insights Into Teacher Reflective Practice During Planning 816
Michael T. Dianovsky, Donald J. Wink

Analyzing Online Knowledge-Building Discourse Using Probabilistic Topic Models 823
Weiyi Sun, Jianwei Zhang, Hui Jin, Siwei Lyu

Developing an Orchestrational Framework for Collective Inquiry in Smart Classrooms: SAIL Smart Space (S3) 831
Mike Tissenbaum, James D. Slotta

Bidirectional Analysis of Creative Processes: A Tool for Researchers 839
Alecia Marie Magnifico, Erica Rosenfeld Halverson, Christopher T. Cutler, TJ Kalaitzidis

Learning from Self-Explaining Emergent Phenomena 847
Kasia Muldner, Winslow Burleson, Michelene T. H. Chi

The Programmers' Collective: Connecting Collaboration and Computation in a High School Scratch Mashup Coding Workshop 855
Deborah A. Fields, Veena Vasudevan, Yasmin B. Kafai

A Tale of Two Worlds: Using Bifocal Modeling to Find and Resolve "Discrepant Events" Between Physical Experiments and Virtual Models in Biology 863
Tamar Fuhrmann, Shima Salehi, Paulo Blikstein

"Deep Hanging": Mentors Learning and Teaching in Practice 871
Déana Scipio

Not a Magic Bullet: The Effect of Scaffolding on Knowledge and Attitudes in Online Simulations 879
Ido Roll, Adriana Briseño, Nikki Yee, Ashley Welsh

Explanations that Make Sense: Accounting for Students' Internal Evaluations of Explanations 887
Shulamit Kapon, Orit Parnafes

Exploring A Digital Tool for Exchanging Ideas During Science Inquiry 895
Camillia Matuk, Marcia C. Linn

Framing Sociocultural Interactions to Design Equitable Learning Environments 903
Bryant Jensen

Learning to Practice Data-Driven Instructional Leadership: Confronting Cultural and Historical Contradictions 911
Raymond Kang

Tensions and Possibilities for Political Work in the Learning Sciences 919
Angela N. Booker, Shirin Vossoughi, Paula K. Hooper

"So, I think I'm a Programmer Now": Developing Connected Learning for Adults in a University Craft Technologies Course 927
Deborah A. Fields, Whitney L. King

Report and Reflection Papers

Metacognitive Planning and Monitoring: 9th Graders Performing a Long-Term Self-Regulated Scientific Inquiry in A Complex System 937
Billie Eilam

Investigating the Effect of Curricular Scaffolds on 3rd-Grade Students' Model-Based Explanations for Hydrologic Cycling 942
Laura Zangori, Cory T. Forbes, Christina V. Schwarz

Exploring How Mobile Technology Provides Inquiry Supports for Middle School Students in Conducting Scientific Practices in a Ubiquitous Learning Context 947
Wan-Tzu Lo, Alex Kuhn, Chris Quintana, Ibrahim Delen, Steven McGee, Jennifer Duck

Empowering Under-Represented Middle School Youth in Engineering Knowledge and Productive Identity Work 952
Angela Calabrese Barton, Daniel Birmingham, Edna Tan, Takumi Sato

Becoming a Youth Worker in a Classroom Community of Practice 957
Laurie Ross

Designing for Engagement in Environmental Science: Becoming "Environmental Citizens" 962
Susan Bobbitt Nolen, Gavin Tierney, Alexandra Goodell, Nathanie Lee, Robert D. Abbott

Student Regulation of Collaborative Learning in Multiple Document Integration 967
Jun Oshima, Ritsuko Oshima, Keita Fujii

Evaluating Lesson Design and Implementation within the ICAP Framework 972
Rod D. Roscoe, Pedro J. Gutierrez, Ruth Wylie, Michelene T. H. Chi

Sequencing Sense-Making and Fluency-Building Support for Connection Making between Multiple Graphical Representations 977
Martina A. Rau, Vincent Aleven, Nikol Rummel

Taking DALITE to the Next Level: What Have We Learned from a Web-Based Peer Instruction Application? 982
Elizabeth S. Charles, Chris Whittaker, Nathaniel Lasry, Michael Dugdale, Kevin Lenton, Sameer Bhatnagar, Jonathan Guillemette

Representational Competence and Spatial Thinking in STEM 987
Mike Stieff, Matthew Lira, Dane DeSutter

Expansive Framing and Preparation for Future Learning in Middle-School Computer Science 992
Shuchi Grover, Roy D. Pea, Stephen Cooper

Conceptualizing Teacher's Practices in Supporting Students' Mathematical Learning in Computer-Directed Learning Environments 997
Aaron M. Kessler, Melissa D. Boston, Mary Kay Stein

Teaching Struggling Middle School Readers to Comprehend Informational Text 1002
Donna Caccamise, Angela Friend, Christine Groneman, Megan Littrell-Baez, Eileen Kintsch

Becoming a Computer Scientist: Early Results of a Near-Peer and Social Justice Program with Latino/a Children 1007
Jill Denner, Jacob Martinez, Heather Thiry

Science, Technology, Body and Personhood: The Concept of Health Emerging in High-Tech Modern Medicine Practice 1012
Federica Raia, Martin Cadeiras, Ali Nsair, Daniel Cruz, Grecia Ramos, Claire Alvarenga, Valeria M. Rivera, Kristina Barrientos, Mario C. Deng

“With-Me-Ness”: A Gaze-Measure for Students’ Attention in MOOCs 1017
Kshitij Sharma, Patrick Jermann, Pierre Dillenbourg

A Design Inquiry: Bridging Assessment and Curriculum Frameworks to Engage Students in Science Practices 1022
Angela Haydel DeBarger, Erika Tate, Yves Beauvineau, Mingyu Feng, Patricia Schank, Tamara Heck, Michelle Williams

Design Principles for Motivating Learning with Digital Badges: Consideration of Contextual Factors of Recognition and Assessment 1027
Cathy Tran, Katerina Schenke, Daniel T. Hickey

What’s Happening in the "Quantified Self" Movement? 1032
Victor R. Lee

The Nature of Student Thinking and Its Implications for the Use of Learning Progressions to Inform Classroom Instruction 1037
Alicia Alonzo, Andrew Elby

Teaching about Confidence Intervals: How Instructors Connect Ideas Using Speech and Gesture 1042
Elise Lockwood, Amelia Yeo, Noelle Crooks, Mitchell J. Nathan, Martha W. Alibali

Studying Students’ Early-Stage Software Design Practices 1047
Brian A. Danielak, William E. J. Doane

Can Scaffolds from Pedagogical Agents Influence Effective Completion of Sub-Goals during Learning with a Multi-Agent Hypermedia-Learning Environment? 1052
Michelle Taub, Roger Azevedo, Nicholas Mudrick, Erika Clodfelter, François Bouchet

Civilian Analogs of Army Tasks: Supporting Pedagogical Storytelling Across Domains 1057
Andrew Gordon, Mark Core, Sin-Hwa Kang, Catherine Wang, Christopher Wienberg

From Playing a Game to Solving an Equation 1062
Shulamit Kapon

Supporting Conceptual Understandings Outdoors: Findings from the Tree Investigators Mobile Project 1067
Heather Toomey Zimmerman, Susan M. Land, Brian J. Seely, Michael R. Mohney, Gi Woong Choi, Lucy Richardson McClain

Analyzing University Students' Participation in the Co-Design of Learning Scenarios 1072
Iolanda Garcia

Authenticity Matters: Youth and Science Participation in Design-Based Learning Environments 1077
Gavin Tierney, Déana Scipio

Personas as a Powerful Methodology to Design Targeted Professional Development Resources 1082
Adrian Madsen, Sarah B. McKagan, Eleanor C. Sayre, Mathew "Sandy" Martinuk, Alexander Bell

Engaging in Educational Design Processes for Sustainable Learning: Learning and Becoming in Practice 1087
Mona Holmqvist Olander, Clas Olander

Supporting Computational Algorithmic Thinking (SCAT): Exploring the Development of Computational Algorithmic Thinking Capabilities in African-American Middle-School Girls 1092
Jakita O. Thomas

Investigating Student Generated Computational Models of Science 1097
Satabdi Basu, Anton Dukeman, John S. Kinnebrew, Gautam Biswas, Pratim Sengupta

Perspectival Computational Thinking for Learning Physics: A Case Study of Collaborative Agent-Based Modeling 1102
Amy Voss Farris, Pratim Sengupta

Children's Use of Inscriptions in Written Arguments About Socioscientific Issues 1107
Sihan Xiao, William A. Sandoval

An Interactional Analysis of Gaze Coordination during Online Collaborative Problem Solving Activities 1112
Murat Perit Cakir, Selin Deniz Uzunosmanoğlu

Multimedia Educative Curriculum Materials: Designing Digital Supports for Learning to Teach Scientific Argumentation 1117
Suzanna Loper, Katherine L. McNeill, Raphaela Peck, Jeremy Price, Jacqueline Barber

How Good Is This Evidence? Students' Epistemic Competence in Evidence Evaluation 1122
Clark A. Chinn, Leah C.-C. Hung, Randi M. Zimmerman, Ravit Golan Duncan

Mathematical Tasks as Boundary Objects in Design-Based Implementation Research 1127
Raymond Johnson, Samuel Severance, Heather Leary, Susan Miller

Evolution of Communities of Learning Practice in Higher Education: Collective Units of Analysis 1132
Filitsa Dingyloudi, Jan-Willem Strijbos

Recasting the Textbook: Student Creation of Interactive Digital History Texts with Primary Source Documents 1137
Molly B. Zielezinski, Paul Franz

Sources of Affect around Interdisciplinary Sense Making 1142
Benjamin D. Geller, Julia S. Gouvea, Vashti Sawtelle, Chandra Turpen

The Impact of Principle-Based Reasoning on Hands-on, Project-Based Learning 1147
Marcelo Worsley, Paulo Blikstein

Emotional Engagement in Agentive Science Learning Environments 1152
Andrew Morozov, Leslie Rupert Herrenkohl, Kari Shutt, Phonraphee Thummaphan, Nancy Vye, Robert D. Abbott, Giovanna Scalone

Filling in the Gaps: Capturing Social Regulation in an Interactive Tabletop Learning Environment 1157
Abigail Evans, Jacob O. Wobbrock

Reverberating Words and GED 2014 Academic Writing Instruction:Reflecting on a Functional Linguistics-Based Approach to Grammar Foregrounding the Social Concept of Identity 1162
Sasha Lotas

Influence of Public Design Critiques on Fifth Graders Collaborative Engineering Design Work 1166
Michelle E. Jordan

Tensions in a Multi-Tiered Research-Practice Partnership 1171
Samuel Severance, Heather Leary, Raymond Johnson

Volume 3

Symposia

Combining Generation and Expository Instruction to Prepare Students to Transfer Big Ideas Across School Topics 1179
Inga Glogger, Lennart Schalk, Claudia Mazziotti, Nicole R. Hallinen, Armin Barth, Ralph Schumacher, Katharina Gaus, Alexander Renkl, Katharina Loibl, Nikol Rummel, Doris B. Chin, Kristen P. Blair, Daniel L. Schwartz, Katherine McEldoon

The Interplay of Domain-Specific and Domain-General Factors in Scientific Reasoning and Argumentation 1189
Frank Fischer, Christof Wecker, Andreas Hetmanek, Jonathan Osborne, Clark A. Chinn, Ravit Golan Duncan, Ronald W. Rinhart, Stephanie A. Siler, David Klahr, William A. Sandoval

Science Sims and Games: Best Design Practices and Fave Flops 1199
Mina C. Johnson-Glenberg, Caroline Savio-Ramos, Katherine K. Perkins, Emily B. Moore, Robb Lindgren, Douglas B. Clark, Corey Brady, Pratim Sengupta, Mario M. Martinez-Garza, Deanne Adams, Stephen Killingsworth, Grant Van Eaton, Matthew Gaydos, Amanda Barany, Kurt Squire, Nathan Holbert

Teachers' Learning about Equitable Practice through Talk with Colleagues 1209
Ilana Seidel Horn, Irene Yoon, Britnie Delinger Kane, Nicole Bannister, Elizabeth A. van Es, Victoria Hand, Ilana Seidel Horn

Motivating and Broadening Participation: Competitions, Contests, Challenges, and Circles for Supporting STEM Learning 1219
Yasmin B. Kafai, Natalie Rusk, Quinn Burke, Chad Mote, Kylie Peppler, Deborah A. Fields, Ricarose Roque, Orkan Telhan, Karen Elinich, Alecia Marie Magnifico

Research and Design of Learning Experiences for Families 1228
Maureen Callanan, Catherine Eberbach, Shelley Goldman, Jennifer Jipson, Amber Levinson, Elyse Litvack, Megan Luce, Lucy Richardson McClain, Sinem Siyahhan, Carrie Tzou, Jessica Umphress, Tanner Vea, Heather Toomey Zimmerman, Philip Bell

Re-Placing the Body in Children's Learning 1237
Katie Headrick Taylor, Rogers Hall, Jasmine Y. Ma, Ananda Marin, Nathan Phillips

Connecting Learning and Becoming: Studying Epistemologies and Identities as Interconnected, Dynamic Systems 1247
Eli Gottlieb, Leslie Rupert Herrenkohl, Stanton Wortham, Catherine Rhodes, Martin J. Packer, Martha Rocío Gonzalez, Anna Sfard

Toward an Argumentative Grammar for Socio-Cultural/Cultural-Historical Activity Approaches to Design Research 1254
Michael Cole, Yrjö Engeström, Annalisa Sannino, Kris D. Gutiérrez, A. Susan Jurow, Martin J. Packer, William R. Penuel, Raymond Johnson, Samuel Severance, Heather Leary, Susan Miller, D. Kevin O'Neill

Mapping the Distribution of Children's Digital Media Practices: Methodological Innovations and Challenges 1264
Reed Stevens, Katie Headrick Taylor, Lori Takeuchi, Elisabeth Hayes, Sinem Siyahhan, Brigid Barron, Amber Levinson, Rosalia Chavez Zarate, Caitlin Kennedy Martin, June H. Lee, Ellen Wartella, Alexis R. Lauricella, William R. Penuel

Becoming Reflective: Designing for Reflection on Physical Performances 1273
Tom Moher, Cynthia Carter Ching, Sara Schaefer, Victor R. Lee, Noel Enyedy, Joshua A. Danish, Paulo Guerra, Alessandro Gnoli, Priscilla Jimenez Pazmino, Brenda A. López Silva, Leilah Lyons, Anthony Perritano, Brian Slattery, Mike Tissenbaum, James D. Slotta, Rebecca Cober, Cresencia Fong, Andee Rubin

Enrollment of Higher Education Students in Professional Knowledge and Practices 1283
Crina Damşa, Hanni Muukkonen, Sten Ludvigsen, Monika Nerland, Minna Lakkala, Auli Toom, Kari Kosonen, Liisa Ilomäki, Lina Markauskaite, Peter Goodyear, Agnieszka Bachfischer

Becoming More Mathematical: New Directions for Describing and Designing for Positive Dispositions Toward Mathematics 1293
Melissa Kumar, Tesha Sengupta-Irving, Noel Enyedy, Melissa Gresalfi, Jennifer Langer-Osuna, Anna Sfard, Kris D. Gutiérrez

Theorizing Learning in the Context of Social Movements 1302
A. Susan Jurow, Ben Kirshner, José Antonio Torralba, Sherine El Taraboulsi, Leah Teeters, Barbara Guidalli, Nosakhere Griffin-EL, Samuel Severance, Molly Shea, Erik Dutilly, Rogers Hall

Concept Formation in Activity 1308
Yrjö Engeström, Annalisa Sannino, Yuri Lapshin, Maria Safronova, Jaakko Virkkunen, Irene Vänninen, Marco Pereira Querol, Rogers Hall

Cyberinfrastructure for Design-Based Research: Toward a Community of Practice for Learning Scientists 1315
Sharon J. Derry, Alan J. Hackbarth, Sadhana Puntambekar, William A. Sandoval, Carlos Gonzalez, Katerine Bielaczyc, Richard Lehrer, Allan Collins

Is the Sum Greater than Its Parts? Reflections on the Agenda of Integrating Analyses of Cognition and Learning 1323
Mariana Levin, Orit Parnafes, Timothy Koschmann, Andrea A. diSessa, Reed Stevens, Rogers Hall, Joshua A. Danish, Noel Enyedy

Reimagining Cultural Forms, Ethnographic Methods and Researcher Responsibilities in Studying Engineering and Science Learning: Honoring and Building on the Work of Margaret Eisenhart 1332
Cory Buxton, Karen L. Tonso, Heidi Carlone, Angela Johnson, Jrène Rahm

When Friends Argue: Investigating Argumentative Learning Processes in Facebook 1342
Dimitra Tsovaltzi, Christa S. C. Asterhan, Christine Greenhow, Raluca Judele, Thomas Puhl, Rakheli Hever, Thor Gibbons, Melissa Menzer, Baruch B. Schwarz, Armin Weinberger, Douglas B. Clark

Enhancing Self-Regulated Learning through Metacognitively-Aware Intelligent Tutoring Systems 1352
Benjamin Goldberg, Robert Sottilare, Ido Roll, Susanne Lajoie, Eric Poitras, Gautam Biswas, James R. Segedy, John S. Kinnebrew, Elaine Stampfer Wiese, Yanjin Long, Vincent Aleven, Kenneth R. Koedinger, Phil Winne

History Learning and Teaching Today: Learning What? Becoming What? By What Practice? 1362
Yifat Ben-David Kolikant, Sarah Pollack, Bradley Fogo, Tsafrir Goldberg, D. Kevin O'Neill, Laura D'Amico, Sheryl Guloy, Avishag (Abby) Reisman

Learning and Thinking in Practice: Complex Systems Thinking "In the Wild" 1372
Izabel Duarte Olson, Ananda Marin, Douglas L. Medin, Alon Hirsh, Sharona T. Levy, Megan Bang, Priya Pugh, Megan McGinty, Uri Wilensky

Learning as Multi-Dimensional Psychological and Cultural Ecological Spaces..... 1382
Carol D. Lee, Shirin Vossoughi, Kris D. Gutiérrez, Na'ilah Suad Nasir, Maxine McKinney de Royston, Barbara Rogoff

Combining Video Games and Constructionist Design to Support Deep Learning in Play 1388
Nathan Holbert, David Weintrop, Uri Wilensky, Pratim Sengupta, Stephen Killingsworth, Kara Krinks, Douglas B. Clark, Corey Brady, Eric Klopfer, R. Benjamin Shapiro, Rosemary S. Russ, Yasmin B. Kafai

Disrupting Learning: Changing Local Practice for Good..... 1396
Jasmine Y. Ma, Charles Munter, Einat Heyd-Metzuyanim, James G. Greeno, Molly Kelton, Rogers Hall, Melissa Gresalfi

Differing Notions of Responsive Teaching across Mathematics and Science: Does the Discipline Matter? 1406
Andrew Elby, Jennifer Richards, Janet Walkoe, Ayush Gupta, Rosemary S. Russ, Melissa J. Luna, Amy Robertson, Janet E. Coffey, Ann R. Edwards, Miriam Sherin, Elizabeth A. van Es

Making the Most Out of It: Maximizing Learners' Benefits from Expert, Peer and Automated Feedback across Domains..... 1416
Astrid Wichmann, Danielle S. McNamara, Markus Bolzer, Jan-Willem Strijbos, Frank Fischer, Moshe Leiba, Alexandra Funk, Nikol Rummel, Michaela Ronen, Olaf Peters, Susanne Narciss, Hermann Körndle, Rod D. Roscoe, Laura K. Varner, Erica L. Snow, Chris Quintana

Leveraging Educative Approaches to STEM Disciplinary and Instructional Practices 1426
Philip Bell, Jeanne Chowning, Elaine Klein, Veronica McGowan, Tana Peterman, Kerri Wingert, Anna Maria Arias, Elizabeth A. Davis, Annemarie S. Palincsar, April Luehmann

Learning and Becoming Through Art-Making: Relationships among tools, phenomena, people, and communities in shaping youth identity development 1436
Noel Enyedy, Joseph L. Polman, Cynthia Graville-Smith, Megan Bang, Beth Warren, Ann Rosebery, Jeff Burke, Fabian Wagmister, Amy Bolling, Taylor Fitz-Gibbon, Erica Rosenfeld Halverson, Na'ilah Suad Nasir

Teacher Facilitation of Whole-Class Discussion in Secondary History Classrooms 1446
Avishag (Abby) Reisman, Chava Shane-Sagiv, Lisa M. Barker, Bradley Fogo, Joseph L. Polman

Synergistic Scaffolding of Technologically-Enhanced STEM Learning in Informal Institutions 1456
Leilah Lyons, Emma Anderson, Michael Carney, Karen Elinich, Robb Lindgren, Michael Tscholl, Chris Quintana, Jessica Roberts, Joyce Wang, Susan Yoon, Iris Tabak

Education for Sustainability and Resilience in a Changing Climate 1466
Sameer Honwad, Ofelia Mangen Sypher, Christopher Hoadley, Armanda Lewis, Kenneth Tamminga, Rose Honey, Daniel C. Edelson

Learning Across Settings: Towards Transformative Trajectories of Practice 1474
Shirin Vossoughi, Meg Escudé, Fan Kong, Elizabeth Mendoza, Molly Shea

Posters

What Do They Do?: Tracing Students' Patterns of Interactions within a Game-Based Intelligent Tutoring System..... 1481
Erica L. Snow, G. Tanner Jackson, Danielle S. McNamara

Implementation Model for Developing Training Measures to Foster Values in an Organization 1483
Sandra Niedermeier, Heinz Mandl

"It's Not as Bad as Using the Toaster All of the Time": Trade-offs in a Scratch Game about Energy Use..... 1485
Gillian Puttick, Amanda Strawhacker, Debra Bernstein, Elisabeth Sylvan

Controlling for Statistical Dependencies in CSCL Using General Estimating Equations 1487
E. Michael Nussbaum, Gwen C. Marchand

Between the Lines: The Role of Curriculum Materials and Teacher Language in Communicating Ideas about Scientific Modeling 1489
Carrie-Anne Sherwood, Carrie Allen Bemis, Savitha Moorthy, Cynthia D'Angelo, Tina Stanford, Christopher Harris

Exploring the Role of Theory of Mind and Executive Functions in Preschool Children's Hypothesis Testing and Revision 1491
Jamie Liberti, Susan Golbeck

Learning to Survive "Home-Free": Compulsory Learning and the Politics of Freight-Hopping Mobility 1493
F. Alvin Pearman II

Sink or Swim: Understanding the Evolution of User Behaviors in an Online Educational Community 1495
Min Yuan, Lei Ye, Mimi Recker

Using Deficient Models as Scaffolds for Learning Engineering Concepts of Tradeoffs and Optimization 1497
Chandan Dasgupta, Tom Moher

Intersections of Science Learning and Language Development within Scientific Argumentation: Implications for English Language Learners 1499
Maria González-Howard, Katherine L. McNeill

Exploring the Use of Elaborative Interrogation in an Introductory Physics Course 1501
Robert Zisk, Elana Resnick, Eugenia Etkina

Fusing a Crosscutting Concept, Science Practice, and a Disciplinary Core Idea in Single Learning Progression 1503
Hayat Hokayem, Amelia Wenk Gotwals

Redefining Engagement and Participation: The Co-Construction of Student Learning Practices 1505
Simona Goldin, Michaela Krug O'Neill, Shweta Naik

Reflective Decision Making within the Discourse of Urban Elementary Engineering Classrooms 1507
Christopher G. Wright, Kristen B. Wendell, Patricia Paugh

Adventure Learning @ the Learning Sciences 1509
Brant G. Miller, R. Justin Hougham, Christopher Cox, Von Walden, Karla Bradley Eitel

Visualizing Three-Dimensional Spatial Relationships in Virtual and Physical Astronomy Environments 1511
Patricia Udomprasert, Alyssa A. Goodman, Philip Sadler, Erin Lotridge, Jonathan Jackson, Ana-Maria Constantin, Zhihui Helen Zhang, Susan Sunbury, Mary Dussault

Upper-Level Physics Students' Perceptions of Physicists 1513
Paul W. Irving, Eleanor C. Sayre

Designing and Validating a Story-Based Socio-Emotional Learning Assessment Instrument 1515
Mitra Fatolapour, Soyeon Hwang, Mario Piergallini, Julie Shim, Steven Dow, Carolyn Penstein Rosé

Detecting Iterative Cycles of Engineering Design from Student Digital Footprints in Computer-Aided Design Software 1517
Zhihui Helen Zhang, Charles Xie, Saeid Nourian

Enculturation: Contemporary Use in the Learning Sciences from a Historical Perspective 1519
Ornit Sagy, Yotam Hod

How do Children Draw, Describe, and Gesture about Motion? 1521
Donna Kotsopoulos, Michelle Cordy, Melanie Langemeyer, Laaraib Khattak

Design Principles for Science Laboratory Instruction Using Augmented Virtuality Technologies 1523
Crystal J. DeJaegher, Jennifer L. Chiu, Jie Chao

The Role of Feedback in Interest Development in an Out-of-School Engineering Setting 1525
Joseph E. Michaelis, Mitchell J. Nathan

Comprehension SEEDING: Providing real-time formative assessment to enhance classroom discussion 1527
Ruth Wylie, Michelene T. H. Chi, Robert Talbot, Erik Dutilly, Susan Trickett, Brandon Helding, Rodney D. Nielsen

Conditions for Successful Learning-by-Teaching: Lessons Learned in Training Prospective Science Teachers 1529
Moseli A. Mafa

Learners' Intuitions about Geology 1531
Lauren Barth-Cohen, Jonathan Shemwell, Daniel Capps

Development of Integrated Physics Identity through Physics Learning Assistant Experience 1533
Eleanor W. Close, Jessica Conn, Hunter G. Close

Assessment Analytics in CSCL: Activity Theory Based Method 1535
Wanli Xing, Robert R. Wadholm, Sean Goggins

What Does "Doing" Science Mean in the Elementary School Classroom? Bruno Latour, Inscriptional Transformations, and a New Look at Children's Interactions with Phenomena 1537
Kathryn Lanouette, Eric Berson, Kathleen E. Metz

Cognitive Ethnographies of Heterogeneous Engineering Design 1539
Carlye Lauff, Joanna Weidler-Lewis, Kevin O'Connor, Daria A. Kotys-Schwartz, Mark E. Rentschler

The Role of Stated Relationships in Detecting Contradictions Between Multiple Representations in Science 1541
Candice Burkett, Susan R. Goldman, M. Anne Britt

Improving Online Collaboration by Fostering Norm-Oriented Content Based Knowledge Awareness 1543
Michail D. Kozlov, Tanja Engelmann, Richard Kolodziej, Roy B. Clariana

Fostering Awareness Content Creation by Self-Determined Regulation 1545
Tanja Engelmann, Katrin König, Michail D. Kozlov

Learning Integrated STEM Using Tangible Agent-Based Modeling 1547
Gokul Krishnan, Pratim Sengupta

Linked Reading and Writing Using Wikilinking: Promoting Knowledge Building within Technology-Enhanced Classroom Learning Communities 1549
Tamar Novik, Dani Ben-Zvi, Yotam Hod

Building A Framework For the Process of Crafting and Using Definitions 1551
Angela Little

Re-grow Your City: A NetLogo Curriculum Unit on Regional Development 1553
Arthur Hjorth, Uri Wilensky

Colors of Nature: Connecting science and arts education to promote STEM-related identity work in middle school girls 1555
Carrie Tzou, Laura Conner, Stephen Pompea, Mareca Guthrie

Reel Science: Identity Development through Filmmaking 1557
Rachel Chaffee

Identifying and Assessing Computational Thinking Practices 1559
Wendy Martin, Karen Brennan, William Tally, Francisco Cervantes

Undergraduate Attitudes Towards Help-Seeking 1561
Iris Howley, Carolyn Penstein Rosé

Becoming a Professional through Virtual Practice 1563
Tabitha McKay, Andrea Cantarero, Vanessa Svihla, Elizabeth Yakes Jimenez, Tim Castillo

Community-Based Engineering and Novice Elementary Teachers' Knowledge of Engineering Practices 1565
Kristen B. Wendell, Tejaswini Dalvi

Using Classroom-Based Authentic Research Experiences to Foster Scientific Thinking and Representational Competence 1567
Kristy L. Halverson, Kari L. Clase

Scaffolding Argumentation Competence: The Shift from First to Second Order Skill Acquisition 1569
Omid Noroozi, Paul A. Kirschner, Harm J. A. Biemans, Martin Mulder

Multiple-Text Processing in Text-Based Scientific Inquiry 1571
Katherine James, Susan R. Goldman, Mon-Lin Monica Ko, Cynthia Greenleaf, Willard Brown

Building a Learning Progression for Chromosome Segregation Using Phenomenographic Variation Theory 1573
Stanley M. Lo, Stephanie Kim, Su Swarat, Gregory J. Light

Tools for Sustained Student Engagement in InterLACE (Interactive Learning and Collaboration Environment) 1575
Chris Teplovs, Leslie Schneider

"Are You 'In' or Are You 'Out'?" Investigating the Factors Affecting Immersion in a Location-Based AR Game for IBSE 1577
Yiannis Georgiou, Eleni A. Kyza

Collaborative Hypothesis-Building Using Immersive Virtual Environments for Ecosystems Science 1579
Shari Metcalf, Amy Kamarainen, Tina Grotzer, Chris Dede

Understanding Data Variability in Ecosystems: Blending MUVE and Mobile Technologies to Support Reasoning with Real World Data 1581
Amy Kamarainen, Shari Metcalf, Tina Grotzer, Chris Dede

Identifying Affordances of 3D Printed Tangible Models for Understanding Core Biological Concepts 1583
Jodi L. Davenport, Matt Silberglitt, Jonathan Boxerman, Arthur J. Olson

Comparison of Specific and Knowledge Integration Automated Guidance for Concept Diagrams in Inquiry Instruction 1585
Kihyun Ryoo, Marcia C. Linn

Elementary Students Becoming Engineers through Practice 1587
Cathy P. Lachapelle, Christine Gentry, Jonathan D. Hertel, Christopher San Antonio-Tunis, Christine M. Cunningham

Local Ground: A Toolkit Supporting Metarepresentational Competence in Data Science 1589
Sarah Van Wart, Tapan Parikh

Make to Relate: Narratives Of, and As, Community Practice 1591
Colin Dixon, Lee Martin

Gesture Enhancement of a Virtual Tutor via Investigating Human Tutor Discursive Strategies: Forms and Functions for Proportions 1593
Virginia J. Flood, Alyse Schneider, Dor Abrahamson

Beyond 'Solve for x': Integrating Equations with Conceptual Understanding 1595
Matthew Lira

Purposeful Learning across Collaborative Educational Spaces 1597
Teresa Cerratto Pargman, Nuno Otero, Marcelo Milrad, Daniel Spikol, Ola Kutsson, Robert Ramberg

Taking a New Perspective on Spatial Representations in STEM 1599
Dane DeSutter, Mike Stieff

The Power of Networks as an Engineering Sophomore 1601
Janet Y. Tsai, Daria A. Kotys-Schwartz, Daniel W. Knight

Getting your Drift: Activity Designs for Grappling with Evolution 1603
Corey Brady, Michael Horn, Uri Wilensky, Aditi Wagh, Arthur Hjorth, Amartya Banerjee

Modeling the Dynamics of Ontological Reasoning in Physics 1605
Lele Mathis, Ayush Gupta

Finding Productivity in Design Task Tinkering 1607
Gina Quan, Ayush Gupta

Making Mathematical Meaning through Robot Enactment of Mathematical Constructs 1609
Scot Sutherland, Tobin White, Jason Huang, Harry Cheng

Predicting Performance Behaviors during Question Generation in a Game-Like Intelligent Tutoring System 1611
Carol Forsyth, Arthur C. Graesser, Borhan Samei, Breya Walker, Philip Pavlik, Jr.

Promoting Science Identification and Learning through Contemporary Scientific Investigations Using Practice-Focused Instruction 1613
Katie Van Horne

Creating Material Representations of Practice at the Boundary of Professional Development and Classroom Practice 1615
Scott McDonald, Jessica Thompson

Examining How Students Make Sense of Slow-Motion Video 1617
Min Yuan, Nam Ju Kim, Joel Drake, Scott Smith, Victor R. Lee

Identity, Digital Learning Environments and Academic Success 1619
Mirlanda E. Prudent

Science Teacher Pedagogical Design Capacity with Technology in an Integrated Teacher Preparation Program 1621
Aaron M. Kessler, Jennifer Cartier

Family Creative Learning: Engaging parents and children as learning partners in creative technology workshops 1623
Ricarose Roque, Natalie Rusk, Luisa Beck, Xiaodi Chen

Teachers Becoming (Temporary) Engineers to Become Better Teachers 1625
Ayesha Livingston, Jamie Collins, Ara Kooser, Vanessa Svihla

Contradictions on the Process of Becoming a Physics Teacher 1627
Glauco Silva, Alberto Villani

Women Becoming Engineers 1629
Joanna Weidler-Lewis

Characterizing Teachers' Analysis of Student Work 1631
Jason Silverman, Valerie Klein, Wesley Shumar, Cheryl Fricchione

Beyond Databases: Librarians in a Project-Based Language Arts Curriculum 1633
Sarah Evans

Characterizing Teachers' Support of Modeling Practices in Science Classrooms 1635
Deborah Peek-Brown, Shawn Stevens, LeeAnn Sutherland, Sung-Youn Choi, Namsoo Shin, Joseph Krajcik

Advancing Epistemological Frame Analysis to Refine Our Understanding of Inquiry Frames in Early Elementary Interviews 1637
Alejandro Andrade-Lotero, Joshua A. Danish

Capturing Qualities of Mathematical Talk via 'Coding And Counting' 1639
Einat Heyd-Metzuyanim, Michal Tabach, Talli Nachlieli, Carolyn Penstein Rosé

Kinecting in Physics: Student Conceptualization of Motion Through Visualization 1641
Janice Anderson, Steven Wall

"These are Facts": Opportunities for and Barriers to Policy Changes that Support Learning 1643
Abigail Stiles, Julie Bryant, Kersti Tyson, Vanessa Svihla

Tug of War: What is it Good For? Measuring Student Inquiry Choices in an Online Science Game 1645
Nicole R. Hallinen, Julius Cheng, Min Chi, Daniel L. Schwartz

Educational Games in the Classroom: Design-Based Research and Methods for Classroom Mediation 1647
Rachel S. Phillips, Theresa Horstman, Carmen Petrick Smith, Christy Ballweber, Noelle Conforti Preszler, Nancy Vye, John Bransford

"What in the world?" Animated Worlds in Multivariable Modeling with Motion Chart Graph Arguments 1649
Jennifer Kahn

Productive Disciplinary Engagement: Examining Negotiation of Group Activity with Multiple Frameworks 1651
Debra Gilbuena, Marja-Liisa Makela, Tuike Iiskala, Simone Volet, Susan Bobbitt Nolen, Milo Koretsky, Marja Vauras

Everyday Life Science and Engineering:Bridging the Gap Between Formal and Informal Learning among Native American Students in the Northwestern United States 1653
Marcie Galbreath, Rose Honey, Sameer Honwad, Anne Kern, Chris Meyer, Laura Laumatia

Distributed Cognition and Gesture: Propagating a Functional System Through Impromptu Teaching 1655
Robert F. Williams, Simon Harrison

Two Systems, Two Stances: A Novel Theoretical Framework for Game-Based Learning 1657
Mario M. Martinez-Garza, Douglas B. Clark

Keeping Up: Shifting Access to Gateway Resources in a Cycling Community of Practice 1659
Joel Drake, Victor R. Lee

Examining Teacher Assignments and Student Work at the Intersection of Content and Practice in Middle School Science 1661
Britte Haugan Cheng, Jeremy Fritts, Tiffany Leones, Bowyee Gong

Examining the Use of Technology: Affordances and Constraints in a Blended Learning Environment 1663
Déana Scipio, Annie Camey Kuo

Demystifying Success in a Summer Bridge Program: Investigating Students' Intrinsic Motivation and Mastery Goals in the Context of a Learning Analytics Intervention 1665
Steven Lonn, Stephen Aguilar, Stephanie D. Teasley

Designing Collaborative Learning Activities for Two Outcomes: Deep Structural Knowledge and Idea Generation 1667
Rachel J. Lam

Promoting Diversity within the Maker Movement in Schools: New Assessments and Preliminary Results 1669
Paulo Blikstein, Vivian Chen, Andrew Martin

Talk as a Window into Collaborative Lesson Design: Designing a Common Rubric in an Elementary School Work Circle 1671
Kimberley Gomez, Nicole Mancevice, Ung-Sang Lee, Jahneille Cunningham

What Is and Who Can Do Science? Supporting Practice-Linked Identities in Science for Racially/Ethnically Underrepresented Youth 1673
Tammie Visintainer

Workshops

Current Research and Practice on Learning Communities: What We Know, What are the Issues, and Where Are We Going? 1677
Katerine Bielaczyc, Dani Ben-Zvi, Yotam Hod

Analytics for Learning and Becoming in Practice 1680
Golnaz Arastoopour, Simon Buckingham Shum, Wesley Collier, Paul A. Kirschner, Simon Knight, David Williamson Shaffer, Alyssa Friend Wise

Social, Motivational and Affective Dimensions of Learning through Social Interaction 1684
Christa S. C. Asterhan, Sherice N. Clarke

Exposing and Assessing Learners' Epistemic Thinking 1686
Maggie Renken, Clark A. Chinn, Penelope Vargas, William A. Sandoval

MOOCShop 2014 1691
Steven Lonn, Christopher Brooks, Zachary A. Pardos, Barry Peddycord III, Emily Schneider, Ido Roll, Ashley Shaw

Mediated Action and Mediated Discourse Analysis: Studyig Learning and Becoming at the Nexus of Practice 1692
Ingrid de Saint-Georges, Kevin O'Connor

Interaction Analysis of Student Teams Enacting the Practices of Collaborative Dynamic Geometry 1697
Gerry Stahl

Writing Competitive Proposals for Programs in NSF's Division of Research in Learning in Formal and Informal Settings 1698
Ellen McCallie, Christopher Hoadley, Michael J. Ford

Tightening Research-Practice Connections: Taking ISLS Findings to Public Debate 1700
Susan McKenney, Kimberley Gomez, Brian J. Reiser

Constructing Assessment Items that Blend Core Ideas and Science Practices 1703
Angela Haydel DeBarger, Joseph Krajcik, Christopher Harris

NAPLeS: Networked Learning in the Learning Sciences 1704
Freydis Vogel, Frank Fischer, Daniel Sommerhoff

Designing for Student Agency and Authority around Issues of Climate Change 1709
Victoria Hand, Leilah Lyons, Chrystalla Mouza, Elizabeth Walsh

The Learning Theater: Designing a Flexible Reconfigurable Space for Ambitious Learning and Teaching on Campus 1709
Gary Natriello, Hui Soo Chae

Research-Practice Partnerships Workshop 1710
William R. Penuel, Philip Bell

Research-Practice Partnership Workshop for Doctoral and Early Career Researchers 1711
Andrew Shouse, Bronwyn Bevan, Kris D. Gutiérrez, Tricia Harding, Ben Kirshner, Matt D. Krehbiel, April Luehmann

Participant Summaries for Special Workshops

Research-Practice Partnership Workshop for Doctoral and Early Career Researchers

Exploring Longitudinal Outcomes and Trajectories of English Language Learners (ELOTE) 1717
Haiwen Chu

Designing Culturally Relevant Technology Based Learning Environments Across Formal and Informal Spaces 1718
Sameer Honwad

Blended Learning, Analytics, and Mathematics Achievement in Washington D.C. Public Schools 1719
June Ahn

Exploring Learning in Out-of-School Settings 1720
Nan Renner

Designing Open Badges for Teaching Educational Technology Skills to Inservice Teachers and Students 1721
Richard E. West

Designing for Persistence after Failure 1722
Cathy Tran

Toward a Multimodal Semiotics of Mathematics Teaching and Learning 1723
Daniel Ginsberg

Hablemos Sobre Enseñanza/Let's Talk about Teaching": Supporting Conversations about Teaching Practice in Chilean Teacher Networks 1724
Florencia Gomez Zaccarelli

Leadership Development for Collaborative Educational Data Use 1725
Emily S. Lin

Making Learning: Leverage Makerspaces as Learning Environments 1726
Breanne K. Litts

Science Teachers Making Sense of Reform: Learning Together in Professional Development in a High-Stakes Environment 1727
Sara Heredia

Early Career Workshop

Digital Media for Learning and Assessment 1731
Jody Clarke-Midura

Research Summary: Simulations and Games for Science Learning 1732
Cynthia D'Angelo

Collaborative Knowledge Practices in Higher Education: An Analysis of Student Learning in Two Undergraduate Programs 1733
Crina Damşa

Social Aspects of Legitimate Peripheral Participation 1734
Julia Eberle

Scaffolding Undergraduates' Argumentative Reasoning 1735
Julia Gressick

Transforming Technological and Scientific Practice: A Research Summary 1736
Katie Headrick Taylor

Searching for the 'Dialogic' within the Dialogue: Temporal Analysis of Teacher Orchestration in Technology Enhanced Elementary School Classrooms 1737
Andrew Joyce-Gibbons

Undergraduate Students' STEM-Based Computing Practices 1738
Alejandra J. Magana

Integrating the Epistemic, Conceptual, and Social Aspects of Science in Elementary School Instruction..... 1739
Eve Manz

Adding a Human Touch to Personalized Digital Learning..... 1740
Amy Ogan

Learning and Becoming in Practice..... 1741
Jimmy Scherrer

Climate Science in Three Part Harmony: Conceptual, Epistemological and Social Learning Gains..... 1742
Asli Sezen-Barrie

Epistemic Thinking and Evidence-based Reasoning with Multiple Information Sources in Health Decision-Making..... 1743
Shira Soffer-Vital

Listening and Learning in Education..... 1744
Kersti Tyson

Supporting Equitable Participation in Socially Relevant Science: Scientific Practice, Partnerships and Technological Innovation 1745
Elizabeth Walsh

Expressive Technologies for Middle School Math and Science 1746
Michelle Hoda Wilkerson-Jerde

Doctoral Consortium

Learning to Teach Elementary Students to Construct Evidence-Based Claims about Natural Phenomena..... 1749
Anna Maria Arias

Attributions and Epistemology in Conversation: How Math Tutors and Students Co-Construct Accounts of Failure and Knowledge..... 1750
David DeLiema

Identity and Collaborative Learning: Examining the Microdynamics of Social Positioning in Mathematical Group Work 1751
Lesley Dookie

Towards a Theoretical Model of Immersion in Inquiry-Based Science Learning..... 1752
Yiannis Georgiou

Co-Constructing Opportunities for Inquiry: A Cross-Context and Development Ethnography of Young Children's Inquiry..... 1753
Danielle Keifert

Learning to Notice: Supporting Students' Meaningful Engagement in Scientific Practices..... 1754
Abraham S. Lo

Collaborative Learning: An Effective Component of Productive Failure? 1755
Claudia Mazziotti

Drawing and Dissection: Improving Understanding in Anatomy 1756
Dimitrios Panagiotopoulos

Have I Become What I Once Beheld? Identity Construction of the Other in Virtual Worlds 1757
Natasha Anne Rappa

The Role of Evidence and Models in Science Education 1758
Ronald W. Rinhart

Exploring Culturally Responsive Computing Education: Learning with Electronic Textiles in an American Indian Community School 1759
Kristin Anne Searle

Play Your Way Into Math: Supporting Prevocational Students in a Computer Game-based Learning Environment 1760
Judith ter Vrugte

Outside the Margins: Examining the Identities, Communities, and Contexts of Youth Who Attend Alternative Schools 1761
Gavin Tierney

Design-Based Learning in a Community-Based Youth Program: Affordances for Youth Learning and Development 1762
Steven Worker

Children's Use of Inscriptions in Argumentation About Socioscientific Issues 1763
Sihan Xiao

Keynote Presentations

Changing Practice

Jean Lave, University of California, Berkeley

Reactor: Rogers Hall, Vanderbilt University

Jean Lave is a social anthropologist with a strong interest in social theory. Much of her ethnographically-based research concentrates on the re-conceiving of learning, learners, and everyday life in terms of social practice. She has published three books on the subject: Understanding Practice (co-authored with S. Chaiklin, 1993); Situated Learning: Legitimate Peripheral Participation (with E. Wenger, 1991); and Cognition in Practice (1988). More recently her work has taken a historical turn with a collaborative, ethnohistorical research project, Producing Families, Trading in History on the British merchant families engaged in the port wine trade in Portugal – (History in Person: Enduring Struggles, Contentious Practice, Intimate Identities 2000, edited with Dorothy Holland). She finished a book on apprenticeship in Liberia and changing research practice (Apprenticeship in Critical Ethnographic Practice) in 2011 and is currently finishing a book of essays, with Brazilian anthropologist Ana Gomes, to accompany and reflect on Situated Learning. She retired from the University of California, Berkeley in 2006.

Rogers Hall is a learning scientist interested the development, learning, and teaching of STEM conceptual practices that are centrally important in scientific and technical work and that appear (in varied form) as topics and resources in school. His research follows these conceptual practices in and out of school, asking how they are organized, develop through time, and can be designed. A central component of this research asks how conceptual practices are learned and change "in the wild" (e.g., ethnographic studies of work groups in field biology, architecture, urban planning, or archeology). Based on comparative analysis in these studies, Hall and colleagues design and study experimental teaching both in conventional classrooms and linked, community settings. Selected publications include "Counter-mapping the neighborhood on bicycles: Mobilizing youth to reimagine the city" (with K. Taylor), "Talk and conceptual change at work" (with I. Horn), "Modalities of engagement in mathematical activity and learning" (with R. Nemirovsky), "How does cognition get distributed? Case studies of making concepts general in technical and scientific work" (with K. Wieckert and K. Wright), and "Conceptual learning" (with J. Greeno). Hall currently serves as Editor in Chief of the journal, *Cognition and Instruction*.

 © ISLS

Research-Practice Partnerships

Megan Bang, University of Washington
Paul Cobb, Vanderbilt University
Kara Jackson, University of Washington
Michael Sorum, Fort Worth ISD

Moderator: Kris Gutiérrez, University of Colorado

Designing with Communities: Transforming Historically Powered Relations in Teaching and Learning

Megan Bang, University of Washington

Megan Bang is an assistant professor of the Learning Sciences and Human Development in Educational Psychology at the University of Washington. She also teaches in the secondary teacher education program. She is the former Director of Education at the American Indian Center (AIC), where she served in this role for 12 years. She is a former pre-school, high-school, and GED teacher, youth worker, and museum educator. Megan's research is focused on improving the well-being and educational opportunities for youth, families and communities historically disadvantaged by education, with a specific focus on Indigenous communities. She investigates the dynamics of culture, learning, and development in and across the multiple contexts of children's lives. She has been centrally focused on understanding and supporting the complexities of learners navigation of multiple meaning systems in science learning environments. She has worked to understand cross-cultural differences in meaning making about the natural world (both aquatic and terrestrial ecosystems) and how learning in places unfolds. Through community-based methodologies Dr. Bang is working to build community capacity to improve and transform teaching and learning, revitalize culture, language and community well-being, and ensure more Indigenous people are engaged in critical research endeavors.

Partnering with School and Districts to Support All Students' Learning

Paul Cobb, Vanderbilt University
Kara Jackson, University of Washington
Michael Sorum, Fort Worth ISD

Paul Cobb is Professor of Mathematics Education at Vanderbilt University. His current research focuses on improving the quality of mathematics teaching and thus student learning on a large scale, and on issues of equity in students' access to significant mathematical ideas. He received the Hans Freudenthal Medal for a cumulative research program over the prior ten years from the International Commission on Mathematics Instruction in 2005, and the Sylvia Scribner Award from the American Educational Research Association in 2010. He is a member of the National Academy of Education.

Kara Jackson is an assistant professor at the University of Washington. Her research focuses on specifying forms of practice that support all learners to participate in rigorous mathematics and how to re-organize educational contexts to support teachers to develop such forms of practice. From 2007-2010, she was a post-doctoral fellow at Vanderbilt University on a project investigating instructional improvement in middle-grades mathematics at scale; she is currently a co-principal investigator on an extension of this study and leads lines of investigation focused on achieving equity in opportunities to learn mathematics and the coordination of professional learning across role groups and contexts. In 2007, she received her doctorate in Education, Culture, and Society with an emphasis in mathematics education at the University of Pennsylvania Graduate School of Education. She taught high-school mathematics in Vanuatu as a Peace Corps volunteer and was a mathematics specialist, supporting both youth and adults, for the Say Yes to Education Foundation in Philadelphia.

Michael Sorum serves as a Deputy Superintendent for the Fort Worth Independent School District (FWISD). He oversees the Divisions of Teaching and Learning, School Leadership, and Student Support Services. Prior to this role, he served as the Chief Academic Officer for the FWISD and the Providence, Rhode Island School Department where he supervised academics , career and technical education, assessment and data quality, secondary academic advisement and the departments for special student populations: special education, ESL and bilingual education, and gifted education. Sorum taught French, Spanish, ESL, and Reading for ten years at the elementary and secondary levels and has served as a campus instructional guide for mathematics and a

curriculum administrator. He holds degrees in political science and romance languages from L'Université d'Aix-en-Provence, and Portland State University and a master's in Administration, Planning, and Social Policy from Harvard University. His doctorate is from Texas Christian University.

I, Thou, and Them: Distributed Memory and Learning

Geoffrey C. Bowker, University of California, Irvine

Introduced by: Kevin O'Connor, University of Colorado

Geoffrey C. Bowker is Professor at the School of Information and Computer Science, University of California at Irvine, where he directs the Evoke Laboratory, which explores new forms of knowledge expression. Recent positions include Professor of and Senior Scholar in Cyberscholarship at the University of Pittsburgh iSchool and Executive Director, Center for Science, Technology and Society, Santa Clara. Together with Leigh Star he wrote *Sorting Things Out: Classification and its Consequences*; his most recent book is *Memory Practices in the Sciences*. He is currently working on big data policy and on scientific cyberinfrastructure; as well as completing a book on social readings of data and databases. More information can be found at: http://ics.uci.edu/~gbowker.

 © ISLS

Approaches to Studying and Modeling Learning Across Setting and Time

Anna Sfard, University of Haifa
Reed Stevens, Northwestern University
Leona Schauble, Vanderbilt University
Beth Warren, TERC
Marianne Wiser, Clark University

Moderator: Jeremy Roschelle, SRI International

Anna Sfard conducts research and teaches in the domain of learning sciences, with particular focus on the relation between thinking and communication. In her research, she aims to contribute to our understanding of human development at large, and of the growth of mathematical thinking in particular. Her work is guided by the assumption that human thinking is a form of communication. Inspired mainly by the work of Wittgenstein and Vygotsky, this non-dualist tenet eventually leads to the conclusion that our communicational activities is the primary source of all things human. Results of her theoretical and empirical research conducted within this communicational (or "commognitive") framework have been summarized in the monograph *Thinking as communicating: Human development, the growth of discourses, and mathematizing* (2008). Her other volumes, edited or co-edited, include *Learning tools: Perspectives on the role of designed artifacts in mathematics learning* (2002), *Learning discourse: discursive approaches to research in mathematics education* (2003), and *Development of Mathematical discourse: Some insights from communicational research* (2012).

Reed Stevens is a Professor of Learning Sciences at Northwestern University. As an ethnographer of everyday experience, Stevens conducts field studies exploring how learning, thinking, and joint action are comparatively organized in range of socio-cultural contexts. A leading goal of these studies is to understand the ways individuals, groups, and standing cultural practices create, organize, and sustain routine and innovative activity and, in particular, how people learn together. In the past two decades he has conducted field studies spanning homes, schools, libraries, professional workplaces, and museums. He draws on understandings generated in these field studies to design and reorganize learning environments and experiences. One current project called FUSE Studios draws on a decade of informal learning studies to rethink STEM as STEAM learning and engagement, using a metaphor of 'leveling up' in video games. (http://vimeo.com/85162569). Other current work includes field studies of young people's everyday experiences using and learning with media, the design and study of a family game to understand and reorganize household energy consumption, and a field study of early career engineers. He was has co-led two NSF Centers, the Center for the Advancement of Engineering Education (CAEE) and the Learning in Informal and Formal Environments Center (LIFE).

Leona Schauble is a cognitive developmental psychologist with research interests in scientific and mathematical reasoning. Shortly after completing her undergraduate degree, she joined the research staff for *Sesame Street* at the Children's Television Workshop. Her subsequent fifteen years at CTW provided practical experience in research and the design of education. In 1987, after completing a PhD in Developmental and Educational Psychology at Columbia University, she went to the Learning Research and Development Center at the University of Pittsburgh as a postdoctoral fellow, where she continued as a Research Scientist until 1992. At the University of Wisconsin and subsequently at Vanderbilt University, she studies learning in both informal and formal educational settings. For example, with The Children's Museum of Indianapolis, the world's largest children's museum, she participated in an NSF-funded project to design and construct an 11,000-foot science gallery that reflects the science knowledge and learning of six- to ten-year-old children. Her current research, conducted in collaboration with Professor Richard Lehrer, is on the origins and development of model-based reasoning in school mathematics and science. In this project, researchers work collaboratively with teachers on an extended basis to generate reform in teaching and learning of mathematics and science, at levels from kindergarten through middle school.

Beth Warren is co-Director of the Chèche Konnen Center at TERC. Prior to joining TERC in 1990, she was Senior Scientist in the Education Group at BBN Laboratories in Cambridge, MA. In her research she focuses on understanding intersections of learning, teaching, disciplinary subject matter, and historically structured inequalities rooted in language, culture, and race. In recent work funded by NSF, the Chèche Konnen Center has been working in collaboration with the Boston Arts Academy, the city's only public high school for the visual and performing arts, to design and develop an artscience model of expansive learning focused on complex, transdisciplinary phenomena such as water and the human microbiome.

Marianne Wiser received a bachelor's degree in oceanography from the University of Liege, Belgium and a Ph.D. from Massachusetts Institute of Technology. She has been at Clark University since 1981. Dr. Wiser studies conceptual change in children, students, and the history of science. Her main topics of research are symbolic development and science learning. Current projects focus on the development of numerical knowledge and number notation in young children; the development of young children's understanding of the nature and function of printed words (pre-reading skills) and how they come to understand the alphabetic nature of our writing system; young children's ability to use models and maps; and young children's conception of matter, weight, and materials. Another topic of research is teaching and learning physics in high school, with special emphasis on microgenetic processes, mental models, parallels with history of science, and the integration of situated cognition approaches with theories of mental representations.

Jeremy Roschelle is Director of the Center for Technology in Learning at SRI International. He co-leads a group of about 80 multidisciplinary researchers who develop educational technologies, conduct learning sciences research and evaluate programs for the National Science Foundation, U.S. Department of Education, Bill and Melinda Gates Foundation, Li Ka Shing Foundation and other government, philanthropic, and industry clients. Within SRI Education, he also leads projects in three lines of work: Community Building, Evaluating Products, and Digital Learning Innovation. Three running themes in his work are democratizing access to advanced mathematics, the study of collaborative learning and appropriate use of advanced or emerging technologies.

Presidential and Invited Sessions

ICLS 2014 Presidential Session: Learning and Assessment of 21st Century Skills

Panel:
Eleni A. Kyza, Cyprus University of Technology
Jan van Aalst, Hong Kong University
Daniel T. Hickey, Indiana University
Alina A. von Davier, Educational Testing Service
Jan-Willem Strijbos, LMU Munich

Moderator:
Cindy E. Hmelo-Silver, Indiana University

The goal of this interactive panel is to examine how the learning sciences can contribute to learning and assessment of 21st century skills. Although it is widely acknowledged that assessment drives learning, the communities that focus on the effectiveness of learning environments and those that focus on valid and reliable assessment co-exist in parallel, rather than in harmony. In addition, across the learning sciences and assessment communities there may be different priorities or "obsessions" when it comes to assessment of 21st century skills. The panelists were each asked to respond to a set of questions which include:

1. What is your focus on learning and assessment of 21st century skills?
2. From a research perspective, what is the biggest challenge?
3. What is your obsession in terms of what is most important for research or practice?
4. Where is the a role for learning sciences in your obsession, if at all?

The panelists in this session bring different perspectives on important questions related to the session theme that we hope will be a springboard to an interactive discussion among panel members and with the audience.

 © ISLS

Research-Practice Partnerships in Communities

Daniel Gallagher, Seattle Public Schools, USA, djgallagher@seattleschools.org
Nichole Pinkard, DePaul University, USA, npinkard@cdm.depaul.edu
Jasmine Alfonso-Gurneau, (formerly of) American Indian Center of Chicago, Oneida and Menominee Nations
Megan Bang, University of Washington, (formerly of) American Indian Center of Chicago, mbang3@uw.edu
Mary Dempsey, (formerly of) Chicago Public Library, USA
Lori Faber, (formerly of) American Indian Center of Chicago, Oneida Nation
Ananda Marin, Northwestern University, (formerly of) American Indian Center of Chicago, Choctaw descent
Mike Sorum, Fort Worth Independent School District, USA

Abstract: There is increasing interest among learning scientists in forging stronger connections to practice through long-term partnerships with schools, districts, community organizations, and networks or coalitions of people focused on improving learning opportunities. In this session, learning scientists and their partners from settings of practice will share their perspectives on forming and maintaining long-term partnerships across the research-practice divide.

 © ISLS

Children Becoming Collaborators

Barbara Rogoff, University of California Santa Cruz, USA, brogoff@ucsc.edu
Rebeca Mejía-Arauz, ITESO Universidad Jesuita de Guadalajara, Mexico, rebmejia@iteso.mx
Lucía Alcalá, Andrew Coppens, Andy Dayton, Angélica López, and Omar Ruvalcaba,
University of California Santa Cruz, USA
Yolanda Corona Caraveo, Universidad Autónoma Metropolitana, Mexico
Maricela Correa-Chávez, Clark University, USA
Kris Gutiérrez, University of Colorado Boulder, USA
Luis Urrieta, University of Texas, USA

Abstract: This poster symposium focuses on cultural aspects of children learning to collaborate with peers, siblings, and adults, and the ways that families and communities support children learning culturally valued ways of collaborating. The posters and the general discussion will contrast collaborating by blending agendas as an ensemble versus dividing turns and resources. Blending agendas as an ensemble is identified as a key feature of Learning by Observing and Pitching-In in family and community endeavors, and seems to be common in many Indigenous-heritage, Mexican, and Mexican immigrant families.

 © ISLS

Teachers as Designers

Joke Voogt, University of Twente, Netherlands, j.m.voogt@utwente.nl
Susan McKenney, Open University/University of Twente, Netherlands, S.E.Mckenney@utwente.nl
Yael Kali, University of Haifa, Israel, yael.kali@edtech.haifa.ac.il
Alain Breleux, McGill University, Canada
Rebecca Cober & Jim Slotta, University of Toronto, Canada
Bat-Sheva Eylon, Weizmann Institute of Science, Israel
Rebecca Itow, Indiana University, USA
Karen Könings, Maastricht University, Netherlands
Therese Laferrière, Université Laval, Canada
Marcia C. Linn, University of California Berkeley, USA
Lina Markauskaite, University of Sydney, Australia
Richard Reeve, Queen's University, Canada
Ornit Sagy, University of Haifa, Israel
Hyo-Jeong So, National Institute of Education, Singapore
Vanessa Svihla, University of New Mexico, USA
Esther Tan, Ludwig-Maximilians-Universität, Germany
Camillia Matuk, University of California, Berkeley

Abstract: Design of (technology-enhanced) learning activities and materials is one fruitful process through which teachers learn and become professionals. To facilitate this process, research is needed to understand how teachers learn through design, how this process may be supported, and how teacher involvement in design partnerships with researchers impacts the quality of the artifacts created, their implementation, and ultimately, student learning. This session speaks to that need by bringing together diverse researchers who, together with practitioners, have explored various affordances of (technology-enhanced learning) design activities for facilitating teaching, learning, curriculum innovation and teacher professional development.

Where are the Learning Sciences in the MOOC Debate?

George Siemens, Athabasca University, Canada, gsiemens@athabascau.ca
Pierre Dillenbourg, Ecole Polytechnique Fedérale de Lausanne (EPFL), Switzerland, pierre.dillenbourg@epfl.ch
Gerhard Fischer, University of Colorado Boulder, USA, gerhard.fischer@colorado.edu
Danielle McNamara, Arizona State University, USA, dsmcnamara@asu.edu
Nikol Rummel, Ruhr-Universität Bochum, Germany, Nikol.Rummel@rub.de

Abstract: This title reflects the crawling skepticism of learning scientists with respect to the hype of MOOCs. Beyond certain individuals, the learning sciences community has been largely indifferent to the MOOC debate. As MOOCs stand today, the activities proposed to students remain quite rudimentary compared to those developed in learning sciences. Let's consider the two mammals of learning sciences, socio-cultural theories and constructivism. The first one encounters some kind of renaissance through the so-called cMOOCs, for 'connectivist' MOOCs, but this renaissance occurs in the shadow of xMOOCs. At a first glance xMOOCs are in contradiction with constructivism. Their most salient part, video lectures, simply perpetuate a pedagogical practice, lecture, that remains ubiquitous in higher education despite empirical evidence of its low effectiveness. Actually, this critique fades out when one looks more closely at what students do in one, the popular high stake MOOCs: they spend around 25% of their time watching videos, versus 75% engaging in the assignments defined by the teacher. Therefore, despite the lecturing component, an xMOOC could afford learners rich problem solving or guided discovery activities that would have comforted Piaget. The bottle does not make the wine. Could the constructivist MOOCs become the agenda of learning sciences? Should learning sciences stay away from MOOCs or, conversely, go on board and attempt to significantly influence their evolution?

Contributions of MOOCs to a Rich Landscape of Learning

Gerhard Fischer, University of Colorado Boulder

Many of the discussions about MOOCS are based on economic perspectives and technology perspectives. Very few contributions analyze MOOCs from a learning sciences perspective and put them into a larger context with other approaches to learning and education. My contribution will postulate that MOOCs represent one approach in a rich landscape of learning. I will use two specific frames of reference ("learning about and learning to be" and "learning when the answer is known and is not known") to analyze the strengths and weaknesses of MOOCs.

Gerhard Fischer is a Professor of Computer Science, a Fellow of the Institute of Cognitive Science, and the Director of the Center for Lifelong Learning and Design (L3D) at the University of Colorado at Boulder. He is a member of the Computer Human Interaction Academy (CHI; 2007), a Fellow of the Association for Computing Machinery (ACM; 2009), and a recipient of the RIGO Award of ACM-SIGDOC (2012). His research has focused on new conceptual frameworks and new media for learning, working, and collaborating, human-computer interaction, and design. His recent work is centered on social creativity, meta-design, cultures of participation, and rich landscapes for learning.

Natural Language Processing: Opening MOOC Doors to Automatic Assessment and Enhanced Collaborative Dialog

Danielle McNamara, Arizona State University

After only a half decade following the dawn of MOOCs, it is clear that large-scale open access education opens up educational opportunities both for educators and learners that were previously beyond reach. Learning sciences has intrinsic interests in enhancing the design of MOOCs as well as capitalizing on their resultant big data. Among the many applications of learning sciences to MOOCS, automated scoring and natural language processing promise to have substantial impacts on the quality of instruction and collaboration. For example, students' written products can be automatically assessed and provided feedback using automated writing assessment. Collaborative interactions through chat rooms, blog posts, and other collaborative tools may be enhanced through algorithms that identify students' needs, interests, or attitudes. Such applications of natural language processing open up doors not only to improve online courses but also to widen our capabilities of developing automated scoring techniques and enhance our understanding of collaborative dialog.

 © ISLS

Danielle S. McNamara is a Professor in the Psychology Department and Senior Scientist in the Learning Sciences Institute at Arizona State University. Her research involves the development and assessment of natural language processing (e.g., Coh-Metrix), game-based, intelligent tutoring systems (e.g., iSTART, Writing Pal), and the use of interactive dialog in automated tutoring systems. The overarching theme of her research is to better understand cognitive and motivational processes involved in reading, writing, memory, and knowledge acquisition and to apply that understanding to educational practice by creating and testing educational technologies. One focus of her work is on developing methods to improve comprehension and writing success for high school students. She has published over 270 articles and other scholarly writings (see soletlab.com).

How Should Learning Sciences Contribute to the Moocs Debate Or How Should It Benefit From It?

Nikol Rummel, Ruhr-Universität Bochum, Germany

As researchers in the Learning Sciences and, particularly, in CSCL, we have in the past decades accumulated ample experience with and evidence for instructional designs that work versus those that do not work. From our findings (i.e., based on scientific evidence), we have deducted educational principles for designing and orchestrating fruitful learning within computer-based and computer-supported environments. Witnessing the explosive development of so-called MOOCs in recent years, my observation is that so far these courses/ learning environments do not yet live up to their potentials, as they are (often) being developed without building on well-established pedagogical knowledge. It is my firm conviction that in order to fully exploit the potential of what is technically possible for learning, developers and designers, as well as educational researchers and educators need to work hand in hand in the making of MOOCs. An area that will need particular attention in this process is the question of how to incorporate and support collaborative elements within MOOCs.

Dr. Nikol Rummel is professor of Educational Psychology at the Institute of Educational Research at the Ruhr-Universität Bochum, Germany, and adjunct professor in the Human-Computer Interaction Institute at Carnegie Mellon University. Dr. Rummel's research interests center around instructional support for learning in computer-supported settings, and especially in computer-supported collaborative learning (CSCL) settings. Dr. Rummel has been or is PI and Co-PI on various research grants funded by different international organizations, such as: the DFG [German Science Foundation], the European Union, the US National Science Foundation (NSF), and the US Institute of Educational Services (IES). Dr. Rummel is associate editor of the Journal of the Learning Sciences, and editorial board member of the International Journal of Computer-Supported Collaborative Learning and of Learning and Instruction. She is elected member of the Board of Directors and co-chair of the CSCL community within the International Society of the Learning Sciences (ISLS), and was program committee co-chair of the 2013 CSCL conference.

Scale as an Opportunity

Pierre Dillenbourg, EPFL, Switzerland

A specific contribution would be to develop pedagogical models that bring the richness of small-scale environment to the scale of MOOCs. So far, scale is perceived as a great opportunity in terms of opening education but as a filter of pedagogical activities. Some learning activities scale well: how much John learns from watching a video or from answering quizzes will be the same whether there are 10 or 10,000 other students watching the same video. At the opposite, some learning activities, manageable with small classes, do not scale up easily. Could learning sciences explore the opposite hypothesis, **"**scale as an opportunity"? An apprentice who works in a company 4 days per week and come to school on the 5th day, discovers the experience of the 19 other apprentices, the variety of responses they encounter to the same problem. Scale can be good. What if they are 10,000 ? Can students share experience, i.e., real shared meaning making, with 10,000 colleagues? Not with current models, but could the learning sciences invent them? I propose a model of MOOCs as weighted geometric graphs that encompass a workflow. The operators of this workflow are technical structures that encompass a pedagogical idea. For instance, an operator that forms teams based on several criteria: if the criterion is the difference of opinions, it will trigger socio-cognitive conflict; if the criterion is the complementarity of knowledge, this will lead to a Jigsaw script. Formalizing pedagogical scripts as graph structures is proposed as a way to run rich learning activities at large scale.

Pierre Dillenbourg is professor of learning technologies at Swiss Federal Institute of Technology in Lausanne (EPFL) and academic director of the Center for Digital Education, whose MOOCs obtained 265'000 registration in one year. Former teacher in elementary school, Pierre graduated in educational science (University of Mons, Belgium). He started to conduct research in learning technologies in 1984. He obtained a

PhD in computer science from the University of Lancaster (UK), in the field of educational applications of artificial intelligence. . His lab, CHILI ("Computer-Human Interaction for Learning and Instruction") explores the frontier between CSCL and HCI, namely novel interfaces for face-to-face collaboration (tangibles, paper computing), robotics, tangible-MOOCs, eye tracking, etc.

Understanding and advancing the MOOC research landscape

George Siemens, Athabasca University, Canada

MOOC hype has produced skepticism amongst academics. The public conversation of MOOCs suggests broad systemic transformation of higher education while largely ignoring the research to date in learning sciences and online learning broadly. This perceived lack of research focus has resulted in growing resistance in many academic quarters to MOOCs as a suitable research domain for serious researchers. In spite of this resistance, research is advancing rapidly, often relying on artificial intelligence, machine learning, and learning analytics. Interestingly, much of this research is being undertaken in communities that have historically not been active in learning sciences. The research being conducted in these communities is not receiving much recognition as it is outside of the distance and online learning, CSCL, and learning sciences communities. This session will provide an overview of the research methods and academic domains that are participating based on a structured analysis of the MOOC Research Initiative submissions (www.moocresearch.com). Through this analysis, it is expected that opportunities will be identified for increased presence of learning sciences researchers to participate in MOOC research.

George Siemens is the Associate Director of the Technology Enhanced Knowledge Research Institute at Athabasca University, and a faculty member in the Centre for Distance Education, and an adjunct in the School of Computing and Information Systems. He has been active in planning, designing and delivering MOOCs since 2008. Recently, he led a Gates Foundation Grant on MOOC Research (funding a total of 28 MOOC grants across different education institutions and MOOC types).

Learning and Becoming through Making and Participatory Media

Julian Sefton-Green, London School of Economics and Political Science, UK, julian@julianseftongreen.net
Kristiina Kumpulainen, University of Helsinki, Finland, kristiina.kumpulainen@helsinki.fi
Karen Brennan, Harvard University, USA, karen_brennan@gse.harvard.edu
Anna Mikkola, University of Helsinki, Finland, anna.mikkola@helsinki.fi
Kylie Peppler, Indiana University, USA, kpeppler@indiana.edu
Elisabeth 'Lissa' Soep, Youth Radio, USA, lissa@youthradio.org

Abstract: Making and maker communities are at the cutting edge of social and economic innovation; participatory media encompasses civic and interest driven ways to create and communicate. This session explores the different potentialities afforded by making and using digital media for young people across a range of learning contexts (formal, out of school, community based and informal). Focusing on dimensions of identity and agency, presentations will question how engaging in digitally mediated expressive and communicative practices offer ways of learning that challenge conventional school pedagogy and curriculum. It will offer analysis of different ways to support and construct learning communities and explore the significance of young peoples' participation in a range of civic and social contexts.

What Does Making and Participatory Media have to offer Learning Science Research?

Whilst there has always been considerable interest in the diverse ways that young people might use forms of out of school culture to create, express, fabricate and communicate in a range of media (see e.g. Barron, 2006), there seems no doubt that access to forms of digital technology and the restructuring of communication ecologies – particularly the World Wide Web – has galvanised an interest in a whole range of media related phenomena – also within the research community of the Learning Sciences.

The terms *'making'* and *'participatory media'* do not accurately describe simple fields of activity but refer in general to a set of practices that have developed within the curriculum; alongside it as part of an extended offer; more organically within discrete cultural practices; and as part of larger social and community-based movements. In general terms we are talking about informal and out of school participation in virtual, asynchronous practices from commercially mediated computer games to more ad hoc alliances of interest driven forums and peer led engagements when we use the term 'participatory media' (see e.g. Sefton-Green & Soep 2007; Kafai, Fields, & Burke, 2010; Peppler, & Kafai, 2010). The term also encompasses opportunities to develop creative and expressive media either individually or as part of new forms of collective social organisation or indeed within more formal traditional frameworks thus allowing young people to play an active part as producers and authors within the wider culture (see e.g. Jenkins *et al* 2007; Kafai, Peppler, & Chapman, 2009; Halverson, 2013).

The idea of *making* derives from an older craft based invention and innovation culture which at his moment, has a particular focus around certain digital processes – in particular programming, coding and devising - as well as using cheap hardware from radio enabled units to 3-D printing to offer new ways of learning about as well as learning through computers and computerisation (Blikstein, 2013; Honey, & Kanter, 2013) Making also includes crucial aspects of design processes and like participatory media offers a host of entrepreneurial and employment focused opportunities (Resnick, & Rosenbaum, 2013). In this invited symposium we conceptualize making in this broadest sense, and thus do not focus merely on STEM learning to which making is often attached.

Both making and engaging with participatory media create possibilities for youth to simultaneously learn about practices and develop their own identities in relation to these practices. We see developing identity at the core of what it means to learn, and constructing practice-linked identities at the core of the process of becoming and re-envisioning oneself in relation to the world (Nasir, & Hand, 2008; Stern, 2008; Wortham, 2006).

Both sets of practices, however diverse and distributed across social and educational contexts, share a number of common features. They both revolve around the development and maintenance of new kinds of learning communities with their own sets of values and norms and which can be independent of traditionally organised institutional educational activities (Hutchins, 1995; Hughes, Jewson, & Unwin, 2007). They both situate young people, traditionally excluded from wider civic participation and/or economic opportunity very much firmly within a continuum of more public facing practices. They both (in principle) have low barriers to access and thus have significant social implications for forms of engagement that are fundamentally equitable in spirit and in practice. They both appear to offer independent, self motivating forms of engagement in complex and challenging activities whilst situating the young person as a real actor in the world and which typically have

been used both as complex learning experiences in her own right or as significant moments in ladders of progression allowing young people to engage in these activities and use them as ways of switching into more conventional academic progression routes (Ito, et al., 2013). Finally, both sets of practices often appear to develop outside of formal educational institutions, they can be self organising and collaboratively structured, they draw on complex sophisticated and current academic knowledge but in grounded and experiential fashions and they can develop forms of expertise and theoretical complexity that has not yet found its expression within traditional academic disciplines: they imply new kinds of knowledge creation and sharing (Sefton-Green 2013).

Their informal institutional basis is both why and how such activities engage marginalised youth and other socially excluded young people: and this theme often is central to accounts of practice (Nasir, & Hand, 2008; Packer, & Goicoechea, 2000). It also underscores why these topics are of great interest to the Learning Science community as they offer sites of learning that are distinct from but clearly related to learning in schools (Walker, & Nocon, 2007).

Because making and participatory media seem to offer forms of learning that have a particular currency at this point in time and can engage a wide constituency of learners from diverse backgrounds these two 'fields' thus appear to offer important sites for investigating types of learning across our society more generally as well as potentially offering models for change and reform within more formally organised learning institutions.

Framing Questions for a Learning Science Enquiry

The Learning sciences research community lacks knowledge about young people's perspective on learning through making or participatory media. At present, research knowledge is scrarce regarding youths' own conceptions of learning from the process of making and engaging in creative/maker cultures. Moreover, while creative/partoicipatory and maker cultures are theoretically open to all, it is unclear why and how they arouse young people's interest and motivation, and potentially broaden participation among diverse youth. There is a need to know which aspects in these cultures motivate and sustain young people's participation and engagement in learning.

To answer these pressing research themes we need to investigate different maker/participatory media cultures and forums – to find out, for example, what difference does it make if making and production are based on competition and/or collaboration, individual and/or collective practices. And what difference these features have on youth engagement, learning and identity given that Learning Sciences researchers are often interested in uncovering and developing the 'design principles' of different learning environments (Brown, 1992). In general, we know little about youth media practice(s) and their role in the organization of youth learning, sociality, and identity formation, either in or out of school.

Research might then be interested in the *quality and particularity of learning relationships in these practices,* how the *contexts of in – and non-formal sites of learning are constituted, regulated and maintained*, and the development of *new forms of disciplinary knowledge* as they are emerging and where they have not yet been sedimented by the academy (Kumpulainen, & Sefton-Green, 2014). More detailed and specifc questions surrounding these issues include:

- How are norms established and maintained in new and marginal making/creative cultures?
- What is the role of 'community' in these learning communities? How is it conceptualised how is it regulated?
- What kind of knowledge is valued and transferred?
- What is the nature of and balance between simulated and 'real-world' activities?
- In what ways does the expressive/participatory/voluntary drive change conventional power relations?
- How informal learning activities are valued, by whom and with what metrics?
- How can such initiatives be scaled and transferred across to other learning domains or should they? This is especially important, since most of this informal activity occurs in homes and outside of organised contexts; how can we be sure that we are inviting youth from non-dominant communities to participate?
- How does the distributed and networked nature of many of these practices give new insight into how we might reformulate and organise curriculum and schooling, especially to support inclusive engagement and learning?
- How do the new tools and systems of meaning making transform our understanding of the relationship between expert, novice and learning progression?
- How and why young people from socially excluded communities find forms of learning that are purposeful and meaningful?
- How youth's practice-linked identities travel across contexts and what consequences this has for their more general engagement with learning?

References

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of Learning Sciences*, *2*(2), 141-178.

Barron, B. (2006). "Interest and Self-Sustained Learning as Catalysts of Development: A Learning Ecology Perspective." *Human Development,* 99, 193–224.

Blikstein, P. (2013). Digital Fabrication and 'Making' in Education: The Democratization of Invention. In J. Walter-Herrmann & C. Büching (Eds.), *FabLabs: Of Machines, Makers and Inventors.* Bielefeld: Transcript Publishers.

Halverson, E. R. (2013). Digital art-making as a representational process. The Journal of the Learning Sciences, 22 (1), 121-162.

Honey, M., & Kanter, D.E. (Eds.) 2013. *Design, Make, Play: Growing the Next Generation of STEM Innovators*. New York: Routledge.

Hutchins, E. (1995). Cognition in the Wild. Cambridge, MA: MIT Press.

Hughes, M., Jewson, N. & Unwin, L. (Eds.) (2007). *Communities of Practice: Critical Perspectives*. Abingdon, England: Routledge.

Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Schor, J., Sefton-Green, J., & Watkins, S.C. (2013). *Connected Learning: An Agenda for Research and Design.* Irvine, CA: Digital Learning and Media Research Hub.

Jenkins, H., Clinton, K., Purushotma, R., Robinson, A., & Weigel, M. (2007). Confronting the Challenges of Participatory Culture: Media Education for the 21st Century. http://digitallearning.macfound.org/atf/cf/%7B7E45C7E0-A3E0-4B89-AC9C-E807E1B0AE4E%7D/JENKINS_WHITE_PAPER.PDF.

Kafai, Y.B., Peppler, K., & Chapman, R. (Eds.) 2009. *The Computer Clubhouse: Creativity and Constructionism in Youth Communities*. New York, NY: Teachers College Press.

Kafai, Y. B., Fields, D. A., & Burke, W. Q. (2010). Entering the clubhouse: Case studies of young programmers joining the online Scratch communities. *Journal of Organizational and End-User Computing, 22*(2), 21-35.

Kumpulainen, K. and J. Sefton-Green (2014). What is Connected Learning and How to Research it? *International Journal of Learning and Media.* In press.

Mcleod, J., & Yates, L. (2006). *Making Modern Lives: Subjectivity, Schooling and Social Change.* Albany: State University of New York Press.

Nasir, N., & Hand, V. (2008). From the Court to the Classroom: Opportunities for Engagement, Learning, and Identity in Basketball and Classroom Mathematics. *Journal of the Learning Sciences, 17*(2), 143-179.

Packer, M. J., & Goicoechea, J. (2000). Sociocultural and Constructivist Theories of Learning: Ontology, Not Just Epistemology. *Educational Psychologist*, 35(4), 227-241.

Petrich, M., Wilkinson, K., & Bevan, B. (2013). It Looks Like Fun but are they Learning? In M. Honey & D. E. Kanter (Eds.), *Design, Make, Play: Growing the Next Generation of STEM Innovators*. London: Routledge.

Resnick, M., & Rosenbaum, E. (2013). Designing For Tinkerability. In M. Honey, and D. E. Kanter (Eds.), *Design, Make, Play: Growing the Next Generation of STEM Innovators* (pp. 163-181). London: Routledge.

Sefton-Green, J., & Soep, L. (2007). Creative media cultures: Making and learning beyond the school. In L. Bresler (Ed.), *International Handbook of Research in Arts*. Amsterdam: Springer

Sefton-Green, J. (2013). *Learning at Not-School: A Review of Study, Theory, and Advocacy for Education in Non-Formal Settings*. Cambridge Mass: MIT Press.

Stern, S. (2008). Producing sites, exploring identities: Youth online authorship. In D. Buckingham (Ed.), *Youth, identity, and digital media* (pp. 95-117). Cambridge, MA: MIT Press.

Walker, D., & Nocon, H. (2007.) Boundary-Crossing Competence: Theoretical Considerations and Educational Design. *Mind, Culture and Activity, 14*(3), 178–195.

Wenger. E. (1998). *Communities of Practice. Learning, Meaning and Identity.* Cambridge: Cambridge University Press.

Wortham, S. (2006). *Learning identity: The joint emergence of social identification and academic learning*. New York: Cambridge University Press.

Papers

Leveling Transparency via Situated Intermediary Learning Objectives (SILOs)

Dor Abrahamson, Kiera Chase, UC Berkeley, 4649 Tolman Hall, Berkeley, CA 94720-1670
dor@berkeley.edu, kiera.chase@berkeley.edu
Vishesh Kumar, Rishika Jain, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
k.vishesh@iitg.ernet.in, rishika.j92@gmail.com

Abstract: When designers set out to create a mathematics learning activity, they have a fair sense of its objectives: students will understand a concept and master relevant procedural skills. In reform-oriented activities, students first engage in concrete situations, wherein they achieve situated, intermediary learning objectives (SILOs), and only then they rearticulate their solutions formally. We define SILOs as heuristics learners devise to accommodate contingencies in an evolving problem space, e.g., monitoring and repairing manipulable structures so that they model with fidelity a source situation. Students achieve SILOs through problem-solving with media, instructors orient toward SILOs via discursive solicitation, and designers articulate SILOs via analyzing implementation data. We describe the emergence of three SILOs in developing the activity Giant Steps for Algebra. Whereas the notion of SILOs emerged spontaneously as a framework to organize a system of practice, i.e. our collaborative design, it aligns with phenomenological theory of knowledge as instrumented action.

When mathematics-education designers set out to create a new learning activity, they bear in mind the activity's *ultimate* pedagogical objective. Reform-oriented designers, however, bear in mind *intermediary* objectives, too, for students' immersive experiences in situated, multimodal, spatial–dynamical activities designed to foster grounded understanding of the ultimate target concepts. Broadly, reform-oriented activities unfold in two steps:

- In Step 1, learners interact with media—physical or virtual materials and ready-made objects—to solve problems that require manipulating, organizing, and/or transforming these media with attention to quantitative relations as well as emerging patterns or principles pertaining to these relations.
- In Step 2, learners are guided to reflect on, and rearticulate their insights using normative semiotic systems, including frames of reference, vocabulary, and symbolic notation and to reenact discovered processes as standard algorithms using the formal representations (Diénès, 1971; Freudenthal, 1983).

This paper focuses primarily on Step 1. Step 1 is of immense importance to the construction of knowledge (Kamii & DeClark, 1985; Piaget & Inhelder, 1969; Thompson, 2013). And yet, we find, educational designers have little, if any, conventional forms, nomenclature, or methodology for articulating Step 1 learning objectives prior to the design process. Perhaps, we submit, this disconcerting lacuna in the design toolkit is related to the ultimate futility of attempting to articulate Step 1 learning objectives prior to building and refining activities and observing people engage with them. Namely, Step 1 objectives emerge only through the design process. Yet this emergent nature of a design's Step 1 objectives, we further submit, should not deter us from eventually defining those objectives. This paper resulted from reflecting on an apparent omission in our own design process: Building a certain design, we kept referring nebulously to a set of latent, contextualized, mathematically oriented, informal ideas we wanted students to discover via engaging in its Step 1 activities. The objective of this paper is to name that unnamed class of ideas and define its role within the design process. We will name this class *situated, intermediary learning objectives* (SILOs) and demonstrate how this ontological innovation lends coherence to a comprehensive, complex, multi-stage process. We hope that, through this paper, fellow designers will join us in "learning and becoming in [design] practice" (the ICLS 2014 theme).

In the remainder of this paper we: overview relevant educational-research literature (Section 1); present Giant Steps for Algebra (Chase & Abrahamson, 2013) (Section 2); explain how three SILOs emerged via developing the design materials and analyzing pilot implementation data and how these SILOs inform our technological redesign (Section 3); and offer implications for theories of knowing and learning (Section 4).

Theoretical Background: Constructing Means for Constructing Meaning

When we design concrete activities for mathematics learning, what are our learning objectives for these activities? These are not quite mathematical learning objectives per se, because they may not be articulated in formal register and might not even involve numerical values. And yet we do eventually form clear ideas for what the students should be discovering about the target concepts through engaging in the concrete activities. In so doing, we implicitly exercise a theoretical view on the relation between the manual and the mental. One such view is ascribed to John Dewey, who characterized conceptual learning as the individual's process of

formalizing their reflection on experience—their guided passage from implicit know-*how* through to articulated know-*that*. Such characterizations of grounded understanding are not only vital for building theories of learning but also bear direct implications for the practice of designing effective learning environments that seek to guide children from informal experience to formal concept. This schematic conceptualization of grounded mathematics learning as an experience-to-concept two-step process cuts across multiple theories and frameworks, including our own.

To begin with, our distinction between situated and general knowledge is a hallmark of Realistic Mathematics Education (RME). Freudenthal (1983), founder of RME, developed a pedagogical methodology based on the principle that children should create their own models of problematic realistic situations. Gravemeijer (1999) elaborates on the function of modeling activities in RME, emphasizing the imperative of letting students' models emerge: "The premise here is that students who work with these models will be encouraged to *(re)invent* the more formal mathematics" (p. 159, original italics). This progress from explorative actions to consistent rules that generalize these actions is theorized more explicitly in RME via the formulation of two related constructs, "model of" and "model for." A "model of" results from modeling a particular situation. A general "model for" eventually emerges from noticing homology across mathematically analogous "models of." Our SILOs (situated, intermediary learning objectives) can be viewed as a checklist detailing structural properties and relations inherent to a "model of." In turn, we deliberately articulate the SILOs in linguistic forms that would also capture general conceptual structures, just as in a "model for."

This ontological relation between actions, objects, and concepts has long fascinated theorists of human activity. For example, *distributed cognition* is a theory of human practice that elucidates relationships among participants to a collective human practice and the artifacts that mediate this collaboration (Clark, 2003). Broadly, the array of tools supporting our cognitive activities—pen and paper, calculator, computer, and so on—are *cognitive artifacts*, that is, artificial tools or devices that carry, elaborate, and report information during problem solving (Norman, 1991). As such, mathematical learning can be theorized as developing psychological structures for regulating the mental activity of distributing quantitative problems over available cognitive artifacts. This effect is dialectical: even as we learn to act and think in new ways as facilitated by these tools, they in turn bear the potential of reifying for our reflection what and how we act and think (Hutchins, 2010). It follows that different material instantiations of one and the same mathematical concept may bear different pedagogical affordances, because their uptake forges different cognitive routes, different neural residue. SILOs articulate this residue pragmatically in terms of the models' structural properties that students learn to monitor.

And yet, this emergence of cognitive structures from mediated actions with external media is not at all guaranteed. A novice might learn to problem-solve using a cognitive artifact that embodies a mathematical function yet without ever understanding this function or how the artifact embodies it. Is this cause for concern? We turn to discuss the psychological construct of *transparency*, which captures relations between, on the one hand, artifacts inherent to a cultural practice and, on the other hand, a social agent's understanding of how features of these artifacts mediate the accomplishment of their objectives. Thus when we say that an artifact is transparent, we refer to the *subjective* relation between a particular agent and the artifact (Meira, 2002).

For educational designers, the notion of transparency suggests a particular framing. Namely, the role of designers can be conceptualized as creating learning tools that learners can render subjectively transparent. In a word, the transparency perspective confers upon educators the role of enabling students to see and learn *how* mathematical artifacts do what they do. For example, in a study of physically distributed problem solving, Martin and Schwartz (2005) found that participants generated more salient and transferable conceptualizations of fractions when using "obdurate" square tiles as opposed to classical pie-shaped manipulatives. Why? From the theoretical lens of transparency, the pie pieces obscured the notion of "whole" precisely because the study participants did not need to assume agency in distributing onto those media their tacit sense of whole—the circle implicitly did that work for them. On the other hand, those students who worked with square tiles were obliged to construct the whole themselves, and that more challenging, agentive experience apparently endured.

Whereas mathematical models per se are often static, such as those fraction squares, they are created through active engagement. Indeed, scholars of embodiment pay close attention to perceptuomotor routines as these relate to conceptual knowledge. In particular, when students operate physically within concretized conceptual domains, design-based researchers attend to how the students carry out spatial–dynamical analogs of formal operations (Antle, 2013). An application of embodiment theory to mathematics education is *embodied design* (Abrahamson, 2009), "a pedagogical framework that seeks to promote grounded learning by creating situations in which students can be guided to negotiate tacit and cultural perspectives on phenomena under inquiry" (Abrahamson, 2013, p. 224). When students participate in embodied-design activities, they solve problems that initially do not bear symbolical notation, do not require calculation, and do not call for quantitative solutions; they call only for qualitative judgments, informal inference, or naïve physical actions.

Embodied designs clearly demarcate the two-step design framework that is thematic to this essay and, as such, underscore the informal nature of Step-1 situated, intermediary learning objectives (SILOs). That is, if we theorize perceptual judgment and motor action as bearing seeds of mathematical concepts, then we need

language for bridging actions and concepts. SILOs articulate subtle elements of learners' informal inferential reasoning about perceptual judgments or motor-action solution strategies that they are to discover and refine.

With the introduction of embodied design, our literature survey shifts from evaluating implications of learning theory for pedagogical design to educational research work dealing directly with the development of design frameworks for grounded mathematical learning.

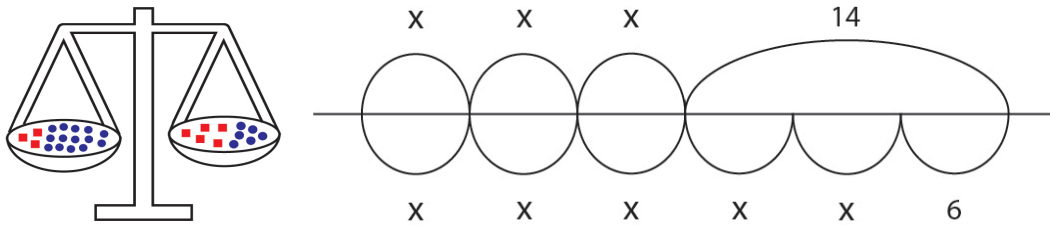
A profound contribution to the design of mathematics learning environments comes from Richard Noss and collaborators, whose learning theory and design frameworks co-emerged dialectically through empirical research studies (Noss, Healy, & Hoyles, 1997). Of particular relevance to our thesis is their set of design heuristics promoting students' *situated abstractions*, "in which abstraction is conceived, not so much as pulling away from context [i.e. the particular features of a situated learning activity], but as a process of constructing mathematical meanings by drawing context into abstraction, populating abstraction with objects and relationships of the setting" (Pratt & Noss, 2010, p. 94, citing Noss & Hoyles, 1996). Pratt and Noss (2010) implicate the epistemological root of mathematical concepts in children's purposeful construction of *utility* for new ideas that are instantiated into designed artifacts in the form of interaction potentialities. The SILOs framework differs from that of situated abstractions in terms of grain size, ontological and epistemological foci, and pedagogical underpinnings. In particular, SILOs articulate a set of initially unavailable interaction constraints that the learner determines, implicates, and wills as potentially conducive to more effective problem solving with a given cognitive artifact; in response, each of these willed constraints is then materialized into the artifact by the instructor who grants the learner's will by enabling into functionality a pre-programmed "hidden" constraint. SILOs are thus functional concretizations of the user's wish-list into working technological features of an interactive device. Yet SILOs are complementary to situated abstractions in the sense that SILOs can be conceptualized as articulating prerequisite structural conditions for enabling and appreciating utility.

In summary, although scholars may differ acutely in their epistemological positions on the constitution of mathematical knowledge, they generally agree that models—forms or structures that learners use in organized activities to promote problem-solving processes—can serve instrumental roles in conceptual development. Having both situated and singled out our proposed heuristic construct of SILO in a legacy of educational theory, philosophy of knowledge, and design frameworks, we now turn to demonstrate this construct's application in an actual case of design practice, namely Giant Steps for Algebra. The next section will explain the design problem that gave rise to this design, and then we explain the design itself.

Setting the Context: Designing Giant Steps for Algebra (GS4A)

The story of learning algebra in schools is often told as the challenge of progressing from arithmetic to algebra. A main character in this story is the "=" sign or, rather, students' evolving meanings for this sign (Herscovics & Linchevski, 1996). When students first encounter algebraic propositions, such as "$3x + 14 = 5x + 6$", their implicit framing of these symbols is *operational*, because the framing will have been fashioned by a history of solving arithmetic problems such as "$3 + 14 =$ __", where you operate on the left-hand expression and then fill in your solution on the right (Carpenter, Franke, & Levi, 2003). Yet algebraic conceptualization of the "=" sign should be *relational*, as this sense contributes to correct treatment of algebraic equations (Knuth, Stephens, McNeil, & Alibali, 2006). Given that the arithmetic visualization of "=" apparently impedes students' transition to algebra, how might this visualization be countervailed? One way is to plant an alternative metaphor.

The *balance metaphor* is undoubtedly the most common visualization of algebraic propositions. This metaphor is typically introduced to students by invoking interactions with relevant cultural artifacts such as the twin-pan balance scale (see Figure 1a). The equivalence-as-balance conceptual metaphor enables a relational, rather than operational, view of algebraic equations. In particular, it grounds the rationale of algebraic algorithms, such as "Remove $3x$ from both sides of the equation," in interactions with a familiar artifact.



a. *b.*

Figure 1. (a) Balance scale and (b) number-line instantiations of "$3x + 14 = 5x + 6$"

Still, students' persistent difficulty in *transitioning* from arithmetic to algebra suggests that the balance metaphor may not be the ideal method for building a relational understanding of equations (Jones, Inglis, Gilmore, & Evans, 2013). Moreover, the historical substitution of twin-pan scales with electronic scales may

have rendered the metaphor unfamiliar to many students. We thus wondered, "What alternative metaphor might facilitate students' passage from arithmetic to algebra?" Our search revealed that Dickinson and Eade (2004) tackled a similar problem. They used the number line as a diagrammatic form for modeling linear equations (see original work in Figure 1b.). Giant Steps for Algebra (GS4A) is based on this "double-measuring-stick" model.

Looking at the number-line diagram in Figure 1b, note the combination of above-the-line and under-the-line symbolic indices of one and the same line segment. This element offers two perceptually contrasting yet conceptually complementary visualizations of a single perceptual stimulus (Abrahamson & Wilensky, 2007). Further note how this number-line diagram "discloses" that $2x + 6 = 14$, so that $2x = 8$, and therefore $x = 4$.

In accord with distributed-cognition theory, this model of algebraic equivalence appears to facilitate the offloading of a rule onto a diagram's inherent logico-figural constraints, so that the problem solver can focus on critical inferences, all the while sustaining a sense of understanding for the solution steps. In the number-line model, but not in the twin-pan model, we are able to construct logical relations between variable and integers directly by attending at a single location to spatial properties such as adjacency and containment.

Finally, inspired by the RME principles, GS4A begins not directly with diagrammatic models of existing symbolic expressions but with an asymbolic situation that the student is required to model diagrammatically. This situation is presented in the form of a narrative about an agent who travels along a path, and the number-line emerges as a "model of" this journey. Per embodied design, we thus sought to engage and leverage students' tacit knowledge about simple ambulatory motion and spatial relations, and per constructivist pedagogy we draw on students' elementary arithmetic fluency.

The GS4A problem narrative depicts a quasi-realistic situation, in which the agent performs two consecutive journeys that begin at the same point of departure and end at the same destination yet differ in process. These two journeys correspond to two equivalent algebraic expressions. For example the algebraic proposition "$3x + 2 = 4x - 1$" is told as a Day-1 journey of "$3x + 2$" and a Day-2 journey of "$4x - 1$", as follows:

> Egbert the Giant has stolen the elves' treasure. He escaped their land and voyaged to a desert island. After docking, Egbert set off walking along a path. You are Eöl the Elf. You are positioned on this island to spy on Egbert and find out what he does with the treasure. Starting from the port and walking straight along the only path, Egbert walked 3 giant steps and then another 2 meters. He buried some of the treasure, covered it up really well, and then went back to the ship, covering up his tracks. On the next day, Egbert wanted to bury more treasure in exactly the same place, but he was not sure where that place was. Setting off along the same path, he walked 4 steps and then, feeling he'd gone too far, he walked back one meter. Yes! He'd found the treasure! He buried the rest of the treasure in exactly the same spot as the day before. Egbert then covered up the treasure as well as all his tracks, so that nobody will know where the treasure is. He returned to the ship and sailed off. Your job is to tell your fellow elves exactly where the treasure is: tell them how many meters they need to walk from the docks to the hidden treasure.

We thus designed GS4A as an environment wherein students develop a notion of variable as a specific quantity: a numerical value that is consistent within a local situation. The specific value of the variable would initially be unknown to the student but could eventually be determined by triangulating available information about the Day-1 and Day-2 journeys. Yet triangulating depictive information—as we learned by tinkering with the design ourselves and observing children attempt to solve the problem—carries certain implicit demands of structural precision and coordination. These "trivial" mechanical details surfaced as conceptually critical.

The Emergence of Situated, Intermediary, Learning Objectives in a Design Process

The GS4A SILOs emerged during our research team's meetings and coalesced over iterated cycles of analyzing empirical data collected in pilot implementations of the design. The SILOs enabled us to coordinate within a single linguistic nexus divergent aspects and objectives of our multi-disciplinary tasks: (1) the target concept (algebra); (2) elements of the design (GS4A); and (3) observations of student behavior (in videotaped studies).

During early trials of the design, we used a variety of different modeling media. This turned out to be fortuitous, in that it ultimately led to us identifying the SILOs. As we argue in Chase and Abrahamson (2013), when the students built a model from scratch, they understood its latent mathematical content better—it was more transparent to them—than in cases where the prefabricated media "did the work" for them (as in Martin & Schwartz, 2005). For example, students were more likely to understand the notion of a variable when they used paper and pencil to painstakingly scale up a drawing that depicted an unfolding sequence of giant steps, than when they were allowed to painlessly stretch an elastic ruler whose intervals scale up uniformly.

Qualitative data analyses suggested the following set of three SILOs for GS4A. (Note that although we articulate the SILOs here as rule-based propositions we do not wish to imply that participants used these forms.)

1. *Consistent measures*. All variable units (giant steps) and all fixed units (meters) are respectively uniform in size both within and between expressions (days);
2. *Equivalent expressions.* The two expressions (Day 1 and Day 2) are of identical magnitude—they share the "start" and the "end" points, so that they subtend precisely the same linear extent (even if the total distances traveled differ between days, e.g. when a giant oversteps and then goes back);
3. *Shared frame of reference*. The variable quantity (giant steps) can be described in terms of the unit quantity (meters).

Articulating the SILOs gradually increased the coherence and effectiveness of our work. In particular, it dawned on us that we should use these SILOs in planning a technological version of our mechanical design. In this technological redesign, the SILOs would form a blueprint for an activity architecture, wherein transitioning from each interaction phase to the next would be linked to demonstrating mastery over one of the SILOs. The idea was thus to step learners through an activity sequence, all the while enabling them to build and sustain subjective transparency of the emerging model. Each SILO would be implemented in this design in the form of some aspect of the model that the learner would be required to build manually (virtually) before that property was instantiated and monitored automatically. Borrowing the notion of "levels" from popular computer games—that is, the gradual rewarding of manifest competency with increased power that is linked to increased task demand—in GS4A we *level transparency.* That is, as the users master each SILO, they receive new control over the environment in the form of enhanced affordances that instantiate that specific SILO automatically.

In GS4A, leveling transparency is engineered as follows. The user encounters a problem narrative and is encouraged to solve it on the screen. A continuous blue path extends horizontally across the screen (see Figure 2). On the left of this line there is a small flag (the "start" location). Below the line there is a standard drawing toolbox with buttons for either selecting a color (giant steps are red, meters are green), toggling between journey days (Days 1 or 2), or editing (removing or clearing model elements). A floating "treasure box" (see in Figure 2, in the top-right corner) can be placed at any location. If a user selects the "Giant Step" button and then clicks on the screen, a red arch will appear that connects the giant's last location along the path (a grey node) to the clicked location (a new grey node). Similarly, "Meters" are green arcs.

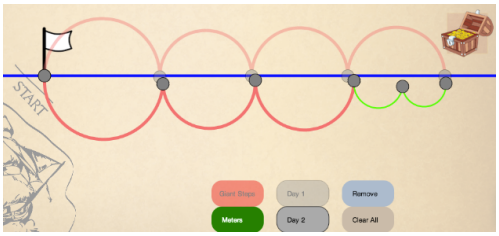


Figure 2. In Level 1, "Free Form," users create all parts of the model manually. Note that the giant steps (red arches) are not quite uniform in size; neither are the meters (green arches).

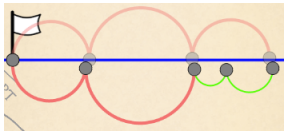
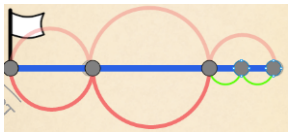
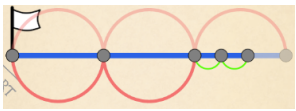
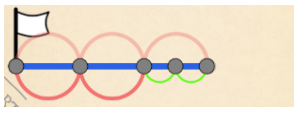
SILOs are psychological constructs—they are about what a child knows (or, at least, about the designer's best understanding of what the child knows). Levels, on the other hand, are technical constructs—they are about an activity's technological affordances, that is, what a pedagogical system performs for you. And yet SILOs and levels are closely related: each SILO articulates a knowledge criterion for entering a new level, and then each level, in turn, orients the child to achieve some next SILO, as outlined in Table 1 (see next page).

We shall now elaborate on this table, referring to its screenshot images. In Level 1, "Free Form," users construct all elements of their model in freehand, analogous to drawing with pencil and paper. Some production imprecision naturally ensues, such as steps that are not quite the same size. The importance of precision (SILO 1) will arise only once the learner attempts to coordinate measures across two journeys, marked above and below the path, and encounters misfits impeding the modeling process. Once users have articulated the imperative of consistency and labored over implementing this aspect in their models, they are evaluated as having graduated SILO 1, "Consistent Measures." As a first concession, the program enters Level 2, "Fixed Meter," in which the system relieves the learner of producing uniform meter units (see also Figure 3, next page).

At this new level, the system supplements manual interaction with optional symbolic interaction. Namely, the learner can now use a control (see in the bottom-right corner of Figure 3) to set a numerical value that determines how many meters will be generated; at a click of a button, the program creates these units as figural elements on the screen. Unburdened by the tedious task of maintaining uniform meters, the user now attempts to equalize the two journeys (Day 1 & Day 2) by adjusting the variable size. Note that one and the same variable, a giant step size, applies both within *each* journey day and across *both* days. As in the case of meters, it is difficult to manually coordinate both within-day and between-days equivalences of variables. Once the user articulates that the variable should be consistent across the entire model, the interface enters Level 3.

In Level 3, "Stretchy," not only is the meter unit size maintained automatically, but the variable size changes uniformly. So when the user drags any of the nodes along the path line, all variable units change size accordingly (please also compare Figures 4a and 4b, two page down). This supplementary affordance enables the user more felicitously to match the end points of Day 1 and Day 2, as follows.

Table 1: Leveling Transparency: Matched SILOs and Levels in Giant Steps for Algebra Technological Design.

SILO	Level	System Constraints, User Activity, and Behavior Criterion	Interface
1. Consistent Measures	1. Free Form	System offers no support in coordinating units or expressions.	
	Activity	User builds all parts of the model manually; is perturbed by units' unequal lengths within and between days; tries to equalize units via small adjustments but witnesses that increasing one unit decreases an adjacent unit sharing a node.	
	Criterion	User expresses frustration in equalizing units.	
2. Equivalent Expressions	2. Fixed Meters	System generates meter units in predetermined size and maintains uniform size automatically.	
	Activity	User builds variables manually; is perturbed by variable units' unequal lengths within/between days; tries to equalize variable units but witnesses that increasing one unit decreases an adjacent unit sharing a node.	
	Criterion	User expresses frustration with managing uniform variable units particularly in an attempt to equalize the two propositions (the spatial extents of Days 1 & 2).	
3. Shared Frame of Reference	3. Stretchy	System monitors for manual adjustment to the size of *any* of the variable units and accordingly adjusts the size of *all* variable units.	
	Activity	User adjusts the variable size to equalize the two propositions	 
	Criterion	User reads off the value of a variable unit in terms of the number of known units (meters) it subtends, e.g., one giant step is 2 meters long.	

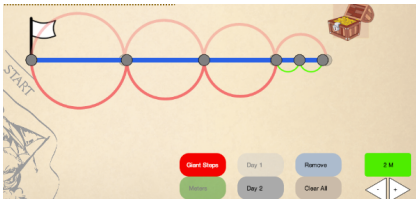
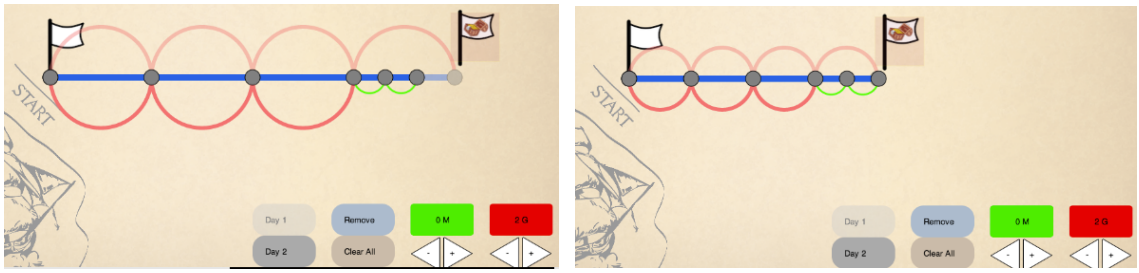


Figure 3. In Level 2, "Fixed Meter," the (green) meters are automatically maintained as uniform in size (and therefore equal to each other), while the variable (red) giant steps are not automatically controlled thus. Users interact with a symbolic control (bottom-right corner) to generate meters.



a. *b.*

Figure 4. In Level 3, "Stretchy," green arches (meters) are invariable and thus equal to each other, while red arches (giant steps) are variable yet always equal to each other via uniform scaling. A new control (bottom-right corner) now enables the user to generate a specified number of giant steps, not only meters.

Note, in Figure 4a, that all the variable units are uniform, both above and below the line path, and yet the two journeys do not end at the same location, so that the only way of aligning the two trips would be by changing the uniform size of the variable (the red arcs). That is precisely what our hypothetical student did, so that the two trips ended in the same location thus determining the value of a single step as 2 meters (see Figure 4b).

A new hypothesis arises from the "leveling transparency" technological design architecture—a hypothesis informing our next study as well as a tentative theoretical insight. Namely, if users were introduced to the activity initially at Level 3, with its full slate of convenient interaction shortcuts, they could not appreciate these features as affordances, because they would not know *what* each feature accomplishes. As such, *learning as constructing transparency is the process of coming to visualize an artifact's features as affordances*.

Closing Words: SILOs Demarcate Structure-Oriented Mathematical Competency

We have introduced a mathematics-education construct we call SILOs—situated intermediary learning objectives. We explained that this construct emerged through our reflective engagement in the process of developing a design for algebra. Sensing the potential of these heuristics as something we might wish to understand, generalize, and share, we reified and refined these tacit elements of our practice in the form of the construct "SILO." SILOs are the structural and logical properties that a learner needs to figure out in order to utilize media made available in a particular learning environment so as to model a particular class of problem situations posed by the activity. Knowing a design's collective SILOs, we maintain, indexes conceptual ontogenesis of a student who is learning target content. Moreover, the creation of a set of SILOs indexes the progress of a designer who is learning about the student's learning process: by articulating the SILOs, the designer comes to know what the students should know who participate in activities enabled by the design. In a sense, SILOs are the educator's heuristics for engineering, orienting, and monitoring the learner's heuristics.

SILOs are not a to-do list of requisite *actions* required by an expert responding to a particular class of problems (e.g., production rules for solving picture-based pre-algebraic problems, Koedinger & Terao, 2002). Rather, SILOs are an artifact's set of necessary *properties*, any of whose violation would elicit from an expert adaptive action. We thus draw on the view of expertise not as the capacity of rote production but rather the skill of responsively recognizing and modifying perceived stimuli so that they embody target structures affording routine practice (Schoenfeld, 1998), such as inferring target information (e.g., the value of x). Of course this modeling skill must be developed. The practical function of SILOs is to organize and coordinate educators' efforts to create, moderate, and evaluate opportunities for learners to reinvent this expertise.

One might be tempted to describe GS4A as an exemplar of technological designs that *scaffold* algebra content. We hesitate to use that common term. In fact, our proposed design architecture for leveling transparency might be described as *reverse scaffolding*. Scaffolding is the asymmetrical social co-enactment of natural or cultural practice, wherein a more able agent performs for novices elements of a complex activity. The novices' participation is thus simplified, so that they experience the activity's purpose, meaning, and efficacy as well as a sense of competence. In GS4A, by way of contrast, the scaffolding is inherent to the design rationale but not the actual activity. That is, the design as a whole is *a fortiori* premeditated to enable and support guided reinvention of a mathematical concept. However, within the environment there is no co-enactment of any steps that students have not yet figured out themselves. The system co-constructs the model only once the student understands the necessity and functionality of each specific property of the model. Thus the pedagogical system relieves users of executing what they *know* to do rather than what they *do not know* to do.

SILOs are subjective achievements—they articulate learners' emergent, idealized system of target relations between reified elements in a problem space; they describe the "things" treated in the situation and imply how to treat them. As such, throughout this manuscript we have spoken of two emergent processes, each of which involves tinkering, discovery, and the objectification of implicit knowledge: (1) the child modeling a situation to infer quantitative information; and (2) the designer modeling the child's behavior to infer learning objectives. These two problem-solving processes are isomorphic, parallel, iterative, and reciprocal.

It is our hope that the idea of SILOs per se as well as the process by which they emerged will resonate with the experiences of fellow designers. A potentially productive focus of such a dialogue would be regarding the ontological status, or pedagogical role, of the external constructions children build as they work on a situated problem, whether concrete or virtual. Additionally, we are fascinated by the *designers'* early process of instantiating mathematical concepts. How does this process transpire? How do designers evaluate the quality, or epistemic fidelity, of these initial conceptual instantiations? We suspect that these two lines of inquiry—about design process and learning process, respectively—will turn out to be more similar than has been formerly suspected and, consequently, mutually informative.

References

Abrahamson, D. (2009). Embodied design: constructing means for constructing meaning. *Educational Studies in Mathematics, 70*(1), 27-47.

Abrahamson, D. (2013). Toward a taxonomy of design genres: fostering mathematical insight via perception-based and action-based experiences. In J. P. Hourcade, E. A. Miller & A. Egeland (Eds.), *Proceedings of IDC 2013* (Vol. "Full Papers", pp. 218-227). New York: The New School & Sesame Workshop.

Abrahamson, D., & Wilensky, U. (2007). Learning axes and bridging tools in a technology-based design for statistics. *International Journal of Computers for Mathematical Learning, 12*(1), 23-55.

Antle, A. N. (2013). Research opportunities: embodied child–computer interaction. *International Journal of Child-Computer Interaction, 1*(1), 30-36.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: integrating arithmetic and algebra in elementary school*. Portsmouth, NH: Heinemann

Chase, K., & Abrahamson, D. (2013). Rethinking transparency: constructing meaning in a physical and digital design for algebra. In J. P. Hourcade, E. A. Miller & A. Egeland (Eds.), *Proceedings of IDC 2013* (Vol. "Short Papers", pp. 475-478). New York: The New School & Sesame Workshop.

Clark, A. (2003). *Natural-born cyborgs: minds, technologies, and the future of human intelligence*. NY: Oxford.

Dickinson, P., & Eade, R. (2004). Using the number line to investigate solving linear equations. *For the Learning of Mathematics, 24*(2), 41-47.

Diénès, Z. P. (1971). An example of the passage from the concrete to the manipulation of formal systems. *Educational Studies in Mathematics, 3*(3/4), 337-352.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Kluwer.

Gravemeijer, K. P. E. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*(2), 155-177.

Herscovics, N., & Linchevski, L. (1996). Crossing the cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics, 30*(2), 39-65.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science, 2*(4), 705-715.

Jones, I., Inglis, M., Gilmore, C., & Evans, R. (2013). Teaching the substitutive conception of the equals sign. *Research in Mathematics Education, 15*(1), 34-49.

Kamii, C. K., & DeClark, G. (1985). *Young children reinvent arithmetic*. NY: Teachers College Press.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? *Journal for Research in Mathematics Education, 37*(4), 297-312.

Koedinger, K. R., & Terao, A. (2002). A cognitive task analysis of using pictures to support pre-algebraic reasoning. In C. D. Schunn & W. Gray (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 542-547). Mahwah: Lawrence.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*(4), 587-625.

Meira, L. (2002). Mathematical representations as systems of notations-in-use. In K. Gravenmeijer, R. Lehrer, B. v. Oers & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 87–104). Dordrecht, The Netherlands: Kluwer.

Norman, D. A. (1991). Cognitive artifacts. In J. M. Carroll (Ed.), *Designing interaction: psychology at the human-computer interface* (pp. 17-38). New York: Cambridge University Press.

Noss, R., Healy, L., & Hoyles, C. (1997). The construction of mathematical meanings: connecting the visual with the symbolic. *Educational Studies in Mathematics, 33*(2), 203-233.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child.* NY: Basic Books (Original work published 1966).

Pratt, D., & Noss, R. (2010). Designing for mathematical abstraction. *International Journal of Computers for Mathematical Learning, 15*(2), 81-97.

Schoenfeld, A. H. (1998). Making pasta and making mathematics: From cookbook procedures to really cooking. In J. G. Greeno & S. V. Goldman (Eds.), *Thinking practice in mathematics and science learning* (pp. 299-319). Mahwah, NJ: LEA.

Thompson, P. W. (2013). In the absence of meaning…. In K. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 57-94). NY: Springer.

Supporting Middle Schoolers' Use of Inquiry Strategies For Discovering Multivariate Relations In Interactive Physics Simulations

Luke D. Conlin, Nicole Hallinen, Daniel L. Schwartz, Stanford University, Graduate School of Education, Wallenberg Hall

Email: lconlin@stanford.edu, hallinen@stanford.edu, danls@stanford.edu

Abstract: Within research on students' inquiry into related variation, several researchers have pointed out the importance of students understanding multiple variable relations. So far, the Control of Variables Strategy (CVS) has demonstrated only limited success in supporting students' discovery of multiple variable relationships. In this report, we present an alternate strategy, which we call the General Principle Strategy (GPS). We report on preliminary results of a classroom study where we taught students in two conditions to use CVS or GPS, respectively, in the context of several physics topics. We find evidence that both strategies help students figure out the multivariable relationship underlying the working of a balance scale, as inferred from associations between their performance on a written posttest and on a computer game-based posttest. Based on these results, GPS shows promise as an effective way of teaching multiple variable relations that underlie a wide variety of physics phenomena.

Introduction

Engaging in scientific inquiry allows students to learn science content while participating in the epistemic practices of science. diSessa (2008) has identified two distinct but complementary modes of inquiry prevalent in the literature, which tap into different aspects of authentic disciplinary practices: inquiry into the meaning of concepts, and inquiry into related variation. The latter involves empirically discovering relations between variables, such as the relation between the range of a projectile and its initial speed.

One line of research on inquiry in science classrooms has identified specific strategies that scientists use to figure out the causal relations between variables and has explored the effectiveness of explicitly teaching these strategies to support students' inquiry (Chen & Klahr, 1999; Ford, 2005; Kuhn, Pease, & Wirkala, 2009). The most prominent inquiry strategy employed in these studies has been the *control of variables strategy* (CVS). CVS involves discovering relations between variables by designing controlled experiments, changing only one variable at a time to make unconfounded comparisons.

Many studies have reported success in teaching CVS to learners in a variety of age groups, who learn to set up unconfounded comparisons and to draw correct inferences from them (Chen & Klahr, 1999; Ford, 2005; Kuhn et al., 2009). These studies have primarily demonstrated the usefulness of CVS for discovering single variable relationships, where only the main effects of a variable are considered, and not its interaction with other variables. For instance, students might use the CVS strategy to find out whether the length of a spring, its width, or how much weight is hung from it affects how far a spring stretches, but do not explore whether the effect of hanging a weight changes based on, say, the width of the spring (Chen & Klahr, 1999; Ford, 2005). Many relations in science involve multiple interacting variables, and as Kuhn (2007) has pointed out, CVS may not be sufficient for unpacking these relationships. So far, there is little evidence that students' learning of CVS helps with their discovery of the relationship between multiple interacting variables (Kuhn et al., 2009; Kuhn, 2007), although this could be due to a lack of instructional supports for extending CVS to handle multiple interacting variables.

Much of the research on students' inquiry into related variation has focused on the hypothetico-deductive approach to science, which is the logic of inference underlying CVS. This approach begins with the formulation of a hypothesis that is then used to deduce observational consequences. Much less work has been done to explore students' use of an equally valid logic of inference, which has played a comparably important role in science: the inductive approach (Shemwell, Chase, & Schwartz, under review). Induction begins with making observations and synthesizing an underlying principle or explanation. While science educators generally recognize the importance of inductively searching for patterns in data, little work has been done to investigate how to support students in conducting such a search in a systematic way.

In this paper, we present an inductive strategy, *general principle strategy* (GPS), which shows promise for supporting students' inquiry into related variation, particularly for discovering the relationships between multiple variables. This strategy has roots in the history and philosophy of science, dating at least back to Bacon (Shemwell, Chase, & Schwartz, under review), and has strong connections to modern accounts of unification and coherence-seeking in science. The general approach involves examining all the data to find one

underlying general explanation. This can apply to a broad array of contexts where CVS may be impracticable (e.g., the historical discovery that the evening star and the morning star were in fact the same object, Venus), but it can also be used as an alternative for making comparisons that establish relations between variables.

GPS offers another way of making unconfounded comparisons. Instead of making pairwise comparisons based on dependent variables (as in CVS), the GPS approach to is to make comparisons across cases based on a common outcome, then to look for common characteristics. The logic of GPS involves using the dependent variable to make inferences about the independent variables, while for CVS the logic of inference proceeds from independent variables to the dependent. For instance, given the top speeds of a set of airplanes with different wing lengths, body shapes, and tail configurations, the GPS approach would be to look at the fastest planes and see what their common characteristics are. The CVS approach would be to pick a characteristic (wing length) and vary only that characteristic to see if the speeds are different.

Other studies have explored ways of supporting students in looking across cases to find a general explanation, for example, by having students invent an index that could apply to multiple contrasting cases. Schwartz, Chase, Oppezzo, & Chin (2011) compared two instructional methods for teaching 7th & 8th grade students the ratio concept underlying density. They found that students who were instructed to invent a "crowdedness" index that could apply to multiple contrasting cases better learned and applied the ratio concept to new physics topics compared with students who were told the ratio concept and given cases to practice. Chi, Dohmen, Shemwell, Chin, Chase, & Schwartz (2012) found improved learning outcomes for undergraduate students who were told to invent a general explanation that can predict the range of several contrasting cases of projectiles. Chase, Shemwell, & Schwartz (2010) compared the general explanation strategy with a Predict, Observe, and Explain (POE) strategy during 50-minute lesson using a physics simulation related to Faraday's law. They added explicit support of the general explanation strategy by providing an example from another domain (buoyancy). They found that students who were guided to seek a general explanation across the cases developed a deeper understanding of the vector component nature of magnetic flux than POE students. In these studies, the supports for students seeking a general explanation were largely embedded in the task, rather than being at the focus of extended, explicit instruction.

Given that several studies have shown that explicit instruction of strategies can improve students' learning and transfer of the strategies (Chen & Klahr, 1999), it is of interest to know whether explicit instruction of GPS could enhance students' learning of the strategy. In what follows, we report on the results of a study in which we taught middle school students either the CVS or GPS strategy in the context of several physics topics, over several weeks. We focus on the results of a posttest item designed to assess their use of an inquiry strategy on a novel physics topic. We report on several interesting associations between their choice of inquiry strategy on this item and their performance on a subsequent computer game-based assessment of their discovery of a multivariable relationship. These results substantiate GPS as a useful strategy for supporting students' inquiry into multivariable relationships.

Methods: Teaching and Assessment

In the present study, we taught four classes of middle school students (132 total) one of two strategies for figuring things out in science (CVS or GPS) during seven 50-minute sessions over a three-week period. Each class was randomly split into two conditions, stratified by class grade and gender. We refer to these conditions as CV or GP to disambiguate them from the strategies. The principle difference between conditions was the strategy they learned for doing inquiry. The CV condition received explicit CVS instruction applied to a variety of physics topics, including projectiles, buoyancy, and collisions, with a focus on learning the content through inquiry. The GP condition received explicit GPS instruction applied to the same sequence of physics topics with the same focus on figuring things out through inquiry. The lessons were taught by two instructors, who each taught 2 classes in each condition to counteract class and teacher effects.

In both conditions, the instruction included a variety of activities such as hands-on explorations, worksheets, and computer simulations of physics phenomena. For example, on the 5th day of instruction, students in both conditions were given simple pendulums (strings with metal washers), and asked to figure out what matters for how quickly a pendulum goes back and forth. In the CV condition, students were encouraged to pick a variable (mass, length, angle, etc.) that might affect how quickly a pendulum will go back and forth, and to test its effect by making comparisons that vary only variable at a time. In the GP condition, the students were also tasked with finding out what affects the pendulum period, but their instructions were to conduct their experiments to find multiple ways to make two pendulum swing at the same rate. Both groups were encouraged to determine the causes of changes in the pendulum period. The sequence of lessons for both groups moved from using inquiry strategies to make causal inferences about single variables to considering multiple variable relations.

After the sequence of lessons, each class took a written posttest that included an item assessing their use of either strategy (CVS or GPS) in the context of a new physics topic, racing different balls down a ramp (see Figure 1). The item presented data on five balls that were rolled down the ramp, including their size,

weight, shape, and the outcome (how long it took to them to roll down the ramp). The item asked the students to: (i) decide which balls they would compare to figure out what makes them go fast (the Ramp Comparisons task), and (ii) use the data to decide what matters for how fast a ball reaches the bottom (the Ramp Conclusions task).

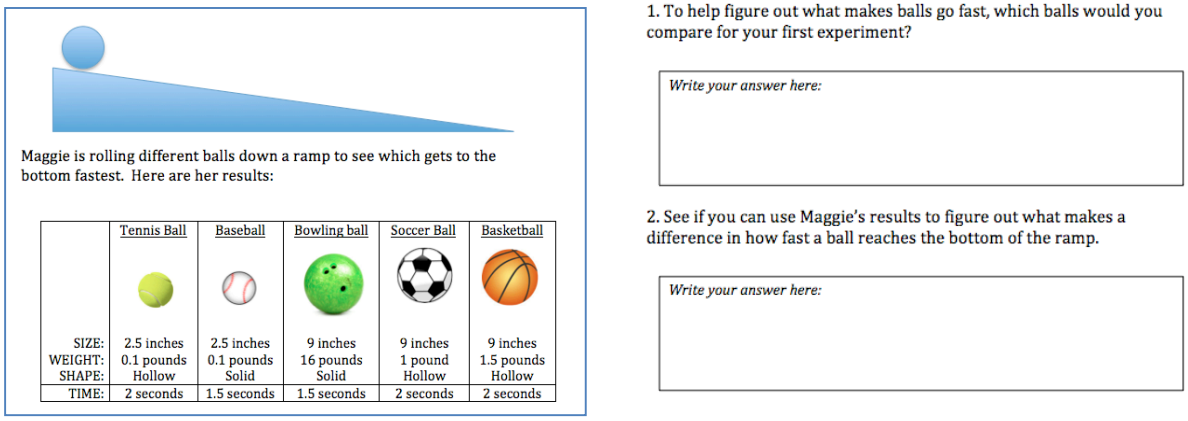


Figure 1. The "Ramp" written posttest item

The day after the written posttest, students took a posttest in the form of a computer game adapted from a physics simulation of a balance scale (Wieman, Adams, & Perkins, 2008). The game included a Challenge mode and an Exploration mode (see Figure 2). In Challenge mode, the students were presented with a sequence of eight challenges: they had to predict whether a given configuration would tip left, tip right, or balance in the middle. In Exploration mode, the students were free to place bricks anywhere on the balance scale and see what happened. Their ultimate goal was to answer eight challenge problems correctly in a row.

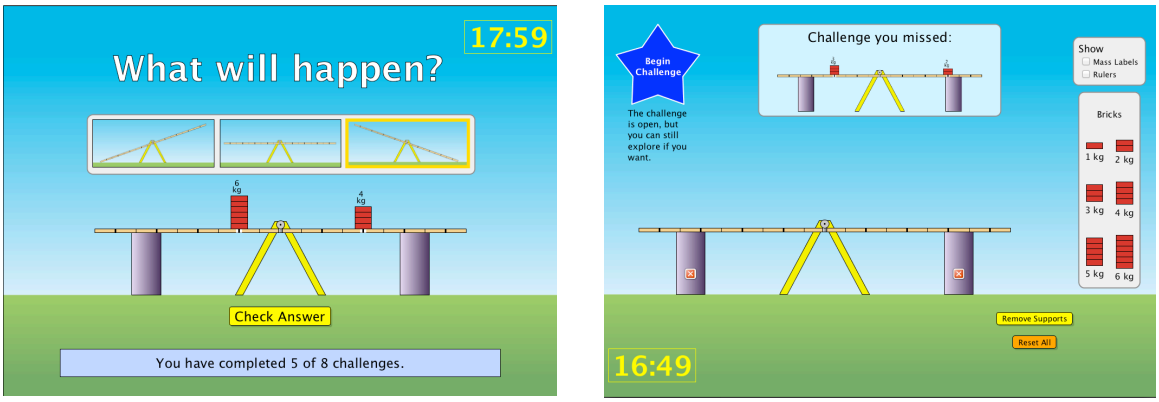


Figure 2. Balance Act posttest Challenge mode (left) and Explore mode (right)

To make predictions about whether the two sides balance, students need to consider multiple variables simultaneously, i.e., the weight on each side and their distances from the fulcrum. The sequence of challenges started off testing just the main effects of each variable (e.g., same amount of bricks on both sides, but farther out on one side) but increased in difficulty to include variable interactions (e.g., one side has more bricks but they are closer to the fulcrum, as in Figure 2). To complete the Challenge mode, the students had to make eight correct predictions in a row. As soon as they got one wrong, they were returned to Exploration mode along with a display of the configuration they missed. They were free to explore, but they could choose to re-enter Challenge mode at any time. When they returned to Challenge mode they had to start again with a whole new set of eight challenges to get through. Performing perfectly in Challenge mode is not likely without figuring out the multiplicative relationship of weight and distance, and so all told, the game serves as an assessment of students' preparedness to learn the multivariate relationship.

Data & Analysis

Of the 132 middle school students, 29 did not return consent forms and were excluded from the analysis, as were 3 students who were absent from either day of posttesting, leaving a sample of n = 100. In what follows, we present an analysis of students' performances on both the written and computer-based posttests. First we

explain the coding scheme for responses to the Ramp Comparisons and the Ramp Conclusions tasks on the written posttest. Then we discuss how students' performance on the Balance Act computer-based assessment game is associated with their strategy use on the Ramp question.

Coding Responses on the Written Posttest

The Ramp item first asked the students to pick which balls to compare in order to figure out what makes a ball go fast down the ramp. There was no specification of how many balls to compare, although most (79%) choose to compare two. We coded the responses to the Ramp Comparisons task as CVS, GPS, or Neither. A student using the Control of Variables Strategy should pick two cases to compare that vary only on one characteristic (weight, shape, or size). There are two possible pairs for which this is the case: tennis ball & baseball or soccer ball & basketball. If the student chose either of these pairs, their response was coded "CVS".

If students are using GPS, they should pick cases that have the same outcome (time down the ramp) then look for what characteristics are common across these cases. In the case of the Ramp Comparisons task there are two ways to pick a common outcome: (i) pick the fastest balls (baseball & bowling ball), which took 1.5 seconds, or (ii) pick the slower balls (tennis ball, soccer ball, & basketball), which took 2 seconds. If a student responded with either of these groupings, their response was coded as GPS (1).

The second question of the Ramp item, which we will refer to as the Ramp Conclusions task, asked students to use the data to decide what affects the time needed for the ball to roll down the ramp. There was no specification of how many factors could be affecting the speed, but most (80%) put only one. Using either strategy should lead to the same (counterintuitive) conclusion for this data, which is "shape" (2). The students' responses were coded as correct if they identified shape (hollow or solid) and did not list any other characteristics.

Two coders independently coded 20% of the responses, which were randomly selected from the posttests. They agreed on 100% of the codes for both Ramp questions before discussion.

Results

Inquiry Strategy on The Ramp Question

Did the students in each condition learn to apply the strategy to the new physics topic? Figure 3 shows a histogram of the strategies used on the Ramp Comparisons task, by condition. Note that the strategy used on this question tends to align with the condition. A Chi-square test of independence shows that this association is significant $\chi^2(2, N = 100) = 6.14$, p < .01. Also note that in the GPS condition (N=53) there were a relatively large proportion of students who used CVS. The converse is not true: in the CVS condition (N=47) only one student used the GPS strategy. This suggests that many students in GP (and by implication, CV) may have already been familiar with CVS. Lastly note that in both conditions (but especially in the GP condition) there is a fairly large proportion of "Neither" codes, i.e., comparisons that did not conform to either strategy.

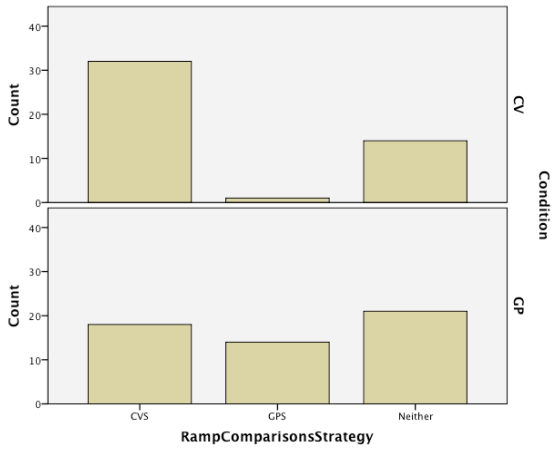


Figure 1. Histogram of inquiry strategies used on the Ramp Comparison task, by condition.

Were the students able to draw the right conclusions from their comparisons? Table 1 is a contingency table for the strategy used on the Ramp Comparisons task with correct responses on the Drawing Conclusions task. There is a significant association $\chi^2(2, N = 100) = 7.33, p < .05$ between using either strategy on the Making Comparisons task and drawing the correct inference in the Drawing Conclusions task. However, Chi-square does not isolate which interaction is driving the effect. To test whether the large proportion of "Neither" codes in GP was behind the association, we collapsed CVS and GPS into a single category "Either" and found

that the association was still significant $\chi^2(1, N = 100) = 7.29, p < .01$. This suggests that students using either strategy were more likely to draw the correct conclusion from their comparison.

The association of drawing the correction conclusion with condition was not significant $\chi^2(1, N = 100) = .17, p > .05$. This could be due to the low incidence rate of people using GPS (n=15), the high number of students in the GP condition using CVS, and the high incidence rate of "Neither" codes. All together, the results suggest that both strategies were helpful for those that used them, but that in future iterations instruction should focus on improving the uptake of GPS.

Table 1: Contingency table for Ramp Comparisons strategy and correct/incorrect Ramp Conclusions.

Ramp Comparisons strategy:	Ramp Conclusion: incorrect	Ramp Conclusion: correct	Total
CVS	12	38	50
GPS	4	11	15
Neither	18	17	35
Total	34	66	100

Strategy Choice and Performance on the Balance Act

To complete the Balance Act Challenge mode, students needed to successfully predict eight challenges in a row. The students were coded for completing the Challenge mode or not. Table 2 is a contingency table for the strategy used on the Ramp Comparisons task with their successful completion of the Balance Act Challenge mode. The association between strategy used on the Ramp Comparisons ramp question and completion of the Balance Act Challenge mode is significant $\chi^2(2, N = 100) = 6.14, p < .05$. The table shows that of the students who used GPS, more than half (60%) completed the Balance Act Challenge. Compare this with students who used the CVS strategy, of which slightly less than half (46%) completed the Balance Act, and with the students who used neither strategy, of which only 26% completed the Balance Act. The table shows that students who used either strategy on the ramp task had a better chance of completing the challenge than those who did not use either strategy.

Table 2: Contingency table for Ramp Comparisons strategy and Balance Act completion.

Ramp Comparisons Strategy	Balance Act challenge not completed	Balance Act challenge completed	Total
CVS	27	23	50
GPS	6	9	15
Neither	26	9	35
Total	59	41	100

We also coded students' performance on the Balance Act according to the maximum number of challenges in a row they got correct. Those who completed the challenge have a max score of 8, and the overall mean was 6.59 (see Figure 4). We find that the strategy used on the Ramp Comparisons task is also associated with the maximum number of challenges won in a row in the Balance Act Challenge mode. An analysis of variance with the maximum score on the balance scale crossed with the inquiry strategy used on the ramp question was significant $F(2,98) = 6.143, p = .003$. The means of the maximum balance scale challenge for those who used either GPS or CVS on the ramp question were significantly higher than those who used neither strategy. The means for GPS are descriptively higher than for CVS, but this difference does not rise to significance $t(63) = 1.08, p = .29$.

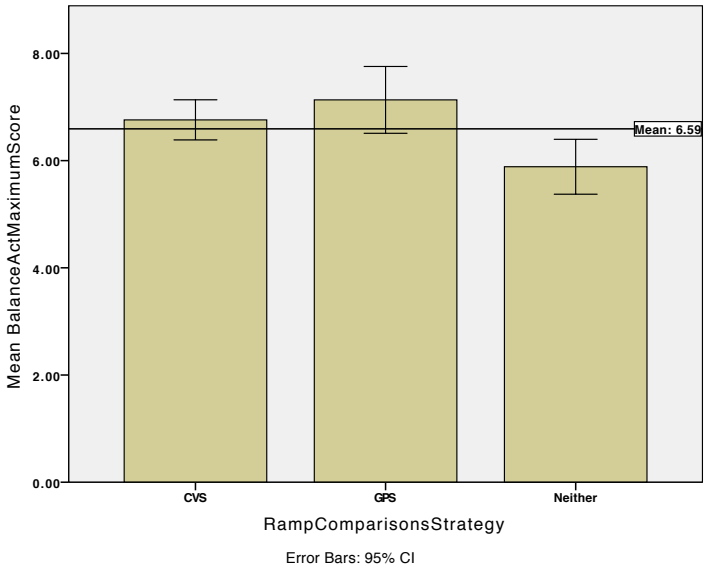


Figure 2. Bar chart comparing means of the maximum number of challenges answered correctly in a row on the Balance Act, by strategy used on the ramp question.

Conclusion & Discussion

The written posttest item assessed whether students learned to apply and draw correct inferences from the instructed inquiry strategy (CVS or GPS) in a new physics context. Analysis shows that the strategies used on this question were associated with each respective condition (CV or GP), and that using either inquiry strategy was significantly associated with finding the correct answer. The computer-based posttest assessed whether students figured out a multiple variable relationship of torque in the context of balancing. Analysis suggests that if they used either strategy on the written posttest, they did significantly better at picking up the multivariate relationship (weight x distance) on the computer-based posttest. This is suggestive that the strategies do help them figure out multiple variable relationships.

In this preliminary study, we found evidence that GPS is at least as effective as CVS for figuring out how multiple variables interact with each other. There is also a hint that GPS helps more than CVS, although due to the low incidence of GPS strategy use, the mean difference did not rise to the level of statistical significance. Overall, this study serves as an existence proof that it is possible to explicitly teach the strategy of seeking a general principle to middle school students in ways that help them figure out multiple variable interactions on their own. This is an important implication for instruction.

Further research is needed to verify that students really are using these strategies as they explore with the Balance Act simulation. This could be corroborated by examining their work on Balance Act either through the session log files and/or by videotaping their screen as they work with the PhET, coding when their moves are consistent with GPS or CVS. The lack of significant associations by condition with Ramp Conclusions and Balance Act measures, due to the low overall incidence rate of students using GPS, as well as the high rates of students using CVS or neither strategy, suggests that future studies should focus on exploring ways of teaching GPS more effectively.

Lastly, future work will focus on finding the productive common ground between students' inquiry into related variation and inquiry into the meaning of concepts. diSessa (2008) has described these modes of inquiry as distinct but complementary, pointing out that few studies have examined their intersection. Ideally, exploring related variation in realistic contexts could inform students' understanding of the meaning of the related contexts, for example, by making sense of the counterintuitive conclusions drawn from their application of inquiry strategies.

Endnotes

(1) There is an ambiguous case: if students chose the soccer ball and the basketball, they could either be controlling variables by making a comparison based on the one characteristic being different (weight), or seeking a general explanation based on the common outcome (2 seconds). We therefore took a conservative approach to coding for GPS; for the slow (2 sec) balls, *all three* had to be compared in order to be coded as a GPS response. If just the basketball/soccer ball pair was selected, it was coded as CVS. It is conservative in that it may be throwing out GPS responses, thereby weakening the correlation between the instructional condition and the strategy used on the ramp

task. Anyone comparing this particular pair is much more likely to be doing so based on the CVS strategy, since that is one of 2 possible correct CVS pairs, while for the GPS strategy that would be one of 4 possible GPS pairs. There were only 5 cases of this, and 3 additional GPS pairs that were excluded based on the conservative coding scheme..

(2) Even though most people intuitively expect the mass and the size of the shapes to matter, it turns out that mass and radius both cancel out of Newton's equations (in the limit of low rolling speeds). Hollow objects have greater rotational inertia, and so cannot speed up as quickly, no matter what the mass or the size is.

References

Chase, C., Shemwell, J., & Schwartz, D. (2010, June). Explaining across contrasting cases for deep understanding in science: An example using interactive simulations. *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1* (pp. 153–160). International Society of the Learning Sciences.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child development*, *70*(5), 1098–1120.

Chi, M., Dohmen, I., Shemwell, J., Chin, D., & Chase, C. (2012, April). Seeing the forest from the trees: A comparison of two instructional models using contrasting cases. *Paper presented at the 2012 American Educational Research Association Annual Conference* (pp. 1–10). Retrieved 3/5/2014, from the AERA Online Paper Repository.

diSessa, A. (2008). A "theory bite" on the meaning of scientific inquiry: A companion to Kuhn and Pease. *Cognition and Instruction 26*(4), 560-566. doi:10.1080/07370000802391760

Ford, M. J. (2005). The game, the pieces, and the players: Generative resources from two instructional portrayals of experimentation. *The Journal of the Learning Sciences*, *14*(4), 449–487.

Kuhn, D. (2007). Reasoning about multiple variables: Control of variables is not the only challenge. *Science Education*, *91*(5), 710–726. doi:10.1002/sce

Kuhn, D., Pease, M., & Wirkala, C. (2009). Coordinating the effects of multiple variables: A skill fundamental to scientific thinking. *Journal of Experimental Child Psychology*, *103*, 268–284. doi:10.1016/j.jecp.2009.01.009

Shemwell, J., Chase, C., & Schwartz, D. (under review). Seeking the general explanation: Revisiting the wisdom of Francis Bacon.

Wieman, C., Adams, W., & Perkins, K. (2008). PhET: Simulations that enhance learning. *Science, 322*(5902), 682-683. doi:10.1126/science.1161948

Acknowledgments

This work was funded by NSF 09-602, and The Hewlett Foundation. The authors would also like to thank Doris Chin, Kristen Blair, Joe Prempeh, Min Chi, and Michael Ford for valuable assistance with the study. Additional thanks go to Kathy Perkins, Michael Dubson, and the PhET team.

Klauer's Inductive Reasoning Training as a Cognitive Apprenticeship Approach for Special-Needs Students

Antonia E. E. Baumeister & Heiner Rindermann
Department of Psychology, Technische Universität Chemnitz, Germany
Email: antonia.baumeister@psychologie.tu-chemnitz.de, heiner.rindermann@psychologie.tu-chemnitz.de

Abstract: The inductive reasoning trainings by K. J. Klauer are characterized as a German cognitive apprenticeship approach for all age groups. The effects of program I were compared across two field studies with special-needs students. The paper-pencil-based training version (N_1 = 34 students; M = 8;9 years) resulted in no improvement of the students' cognitive (CFT 1: d_{mean} = -0.17; $\beta_{meanTrain\text{-}Int}$ = -.08) and verbal abilities (HSET: d_{mean} = -0.02, $\beta_{meanTrain\text{-}Verb}$ = .04). However, the PC version (N_2 = 27; M = 7;8 years) resulted in a small positive effect on intelligence (CFT 1: d_{mean} = 0.28; $\beta_{meanTrain\text{-}Int}$ = .13) and verbal performance (HSET: d_{mean} = 0.25, $\beta_{meanTrain\text{-}Verb}$ = .12). Compared to childhood cognitive development, even the small training effects observed in Study 2 are practically meaningful because they correspond to more than half a year of schooling and maturation. Integrating cognitive trainings in preschool to lifelong learning programs is recommended.

Impact of Cognitive Trainings

Cognitive competence is the ability to think, to possess relevant and accurate knowledge and to apply this knowledge intelligently. This competence is required in order to solve complex problems successfully which can range from language learning to other academic or professional and everyday tasks (e.g., Rindermann, Flores-Mendoza, & Mansur-Alves, 2010; Rindermann, Michou, & Thompson, 2011). One aspect of cognitive competence is reasoning, that is, identifying rules and transferring these rules to new problems. According to Klauer and Phye (2008), inductive reasoning consists of detecting regularities and irregularities by finding out similarities and differences of attributes and relations with differently coded contents (e.g., verbal, pictorial, numerical).

Thinking abilities can be improved by different kinds of interventions, for example, by means of educational methods like trainings (e.g., Irwing, Hamza, Khaleefa & Lynn, 2008; Jaeggi, Buschkuehl, Jonides & Shah, 2011; Schmiedek, Lövdén, & Lindenberger, 2010), early education and preschool (e.g., Campbell & Ramey, 1994) or schooling in general (Ceci, 1991). Especially Klauer has developed cognitive trainings for several age groups (e.g. Klauer, 1989, 1991, 2008), and he has empirically investigated their effectiveness. In numerous studies applying individual, dyadic or small group settings to whole classes, positive effects of the training programs were found on academic learning, concentration and intelligence (e.g. Marx, 2005; Sonntag, 2004). The effectiveness of the programs for children and teenagers across 78 evaluation studies was on average d = 0.52 regarding intelligence gains, and d = 0.69 on average across 38 studies regarding academic learning (corrected for dependencies within samples, respectively; Klauer & Phye, 2008). For example, for children with special needs, positive effects were found which were beyond coaching effects and which turned out to be stable in the long-term (e.g., Marx, 2005).

Assuming that positive effects can spread, even a relatively short training could result in positive long-term effects (cf. Klauer & Phye, 2008). Previous studies on the cognitive training with children and teenagers with special needs showed that the positive effects transfer to further areas, for example, language acquisition, reading literacy or academic learning (e.g., Marx, 2006; Sonntag, 2006).

Although previous studies from different groups of researchers have illustrated the effectiveness of Klauer's cognitive trainings with special-needs students, these studies have researched quite homogeneous samples, for example, regarding competence level or regarding clinical pictures (e.g., senior classes of a special school; Sonntag, 2004). However, children attending special education centers mostly have rather different needs. Therefore, studies are needed for answering the research question whether a cognitive training has positive effects with a quite heterogeneous group of special-needs students.

Klauer's Inductive Reasoning Training as a Cognitive Apprenticeship Approach

Cognitive apprenticeship is an instructional approach which aims at fostering not only domain-specific prior knowledge but also general and transferable problem solving strategies such as learning strategies, control strategies, and heuristics as well as the positive experience of intrinsically motivating, cooperative learning tasks (Collins, Brown, & Newman, 1989). The role of teachers is to demonstrate how tasks can be solved, to explain their solution steps and express their thoughts as experts in this process, to structure suitable exercises for their learners, to help them by providing feedback or further information in case of problems, and to reduce guidance gradually in the course of learners becoming more experienced and independent. These processes can be

summarized by the often-cited keywords of *modeling – coaching – scaffolding – fading* (Collins et al., 1989). The role of learners is to discuss about their knowledge (*articulation*), to compare their approach with other learners' approach (*reflection*), and to explore new problems and strategies (*exploration*; Collins et al., 1989).

This approach has proven to be highly effective for improving students' reading comprehension (Palincsar & Brown, 1984), for acquiring mathematical problem solving strategies (Schoenfeld, 1985), for teaching causal reasoning (Hendricks, 2001), for developing writing skills (Bereiter & Scardamalia, 1987), and for knowledge building in general (Scardamalia & Bereiter, 2006). On a conceptual level, as teaching of inductive reasoning, and on the methodological level, Karl Josef Klauer's cognitive trainings can also be characterized as a cognitive apprenticeship approach. The following steps are common to all of his training programs developed for all age groups from preschool till seniority (e.g. Klauer, 1989):

(1) *Exposition:* This is the starting point of each cognitive training program. Klauer (1989) recommends to present authentic tasks in order to illustrate the relevance of inductive reasoning for everyday life. In this phase, the first tasks are solved by analyzing the initial variables, formulating an aim and a plan how to reach it, testing one's hypotheses, and self-reflecting upon the results.

(2) *Development:* The function of this stage is to understand the general structure of inductive reasoning tasks which is to identify similarities or differences between features, to develop and test rules, to describe and compare task categories, to articulate their features, to learn to use the concepts correctly (e.g. 'features', 'relations', 'sameness', 'differentness' etc.), and to link the task categories with solution strategies.

(3) *Application:* This stage serves four functions. First, through practicing to solve many tasks, the skills and strategies are deepened. Second, learners practice to identify task categories upon being confronted with new tasks. Third, a transfer of prior knowledge to new tasks is achieved. Fourth, all skills should be automatized so that tasks can be solved quickly and effortlessly.

Klauer (1989) recommends the following training methods which can, like the abovementioned training steps, clearly be classified into a cognitive apprenticeship approach:

Guided discovery. The learners mainly work in a self-regulated manner during exploring the task classes with their specific features. Only in case of problems, the trainer gives hints by asking helpful questions and correcting errors. Special-needs students may need more guidance than other learners. The aim is to make the learners think aloud or explain to the trainer as much as possible. Core aspects should be summarized and repeated, for example, solution procedures and control strategies. Klauer (1989) emphasizes the importance of developing and using a common language (O'Donnell & O'Kelly, 1994; Rogoff et al., 2003; Vygotsky, 1978) which supports memorization of the learned strategies.

Verbalization and *self-reflection*. The aim of this training method is that learners should think aloud and justify their task solutions. This should induce a more analytic strategy (than mere guessing), activate and strengthen control strategies, and reduce cognitive load. Self-reflection can be supported by asking the learner why he or she made a specific step etc. This method is more effective in dyadic compared to individual settings because peers can learn from each other by mutually adopting the partner's strategies (Lou, Abrami, & d'Apollonia, 2001).

Verbal self-instructions. This method is especially recommended with special-needs students (Klauer, 1989, p. 99) and has proven to be highly effective (e.g. Masendorf & Klauer, 1987). The trainer first models the task solution by making his or her thinking visible. Further, the trainer shows how he or she reacts to own errors and corrects them. For example, he or she can compliment himself or herself for working accurately. In the second step, the learner should instruct the trainer to solve a task by using the directions he or she has heard before. In the third step, the learner uses self-instructions by verbalizing them aloud, then whispery, and finally, by means of inner speech (Vygotsky, 1978).

Thus, external regulation decreases and self-regulation increases in the order of the training methods: verbal self-instructions – verbalization and self-reflection – guided discovery (Klauer, 1989). Depending on the age of the target group, Klauer recommends different forms of a social training. For example, in kindergarten age, he prefers individual trainings with a maximum duration of 20 minutes (Klauer, 1989, p. 111). Dyadic training settings can be applied from the age of primary school; however, the trainer needs to ensure that the learners alternate during task solution so that both learners receive enough practice. Reciprocal teaching (Palincsar & Brown, 1984) can be used by letting one learner self-comment on his or her strategies and the other learner check and question the strategies applied. After each phase, the roles are changed. This is also possible in small groups of 3-4 learners. Klauer recommends learning in homogeneous groups because this supports achievement motivation and prevents competence threats by more knowledgable learners (Butera, Caverni, & Rossi, 2005). Further, a suitable procedure could be to use direct instruction in groups during the exposition phase, small groups during the development phase, and individual training during the application phase (Klauer, 1989).

Further commonalities of all training programs are the six task categories of inductive reasoning (e.g. Klauer, 1989; Lenhard, Lenhard, & Klauer, 2012):

(1) Generalization: These tasks require identifying that different objects share at least one common feature (e.g. what is the commonality of a butterfly, a kite, and a helicopter?).

(2) Discrimination: The differences between features need to be found out (e.g. pick the odd one out: spade – watering can – telephone – garden hose).

(3) Cross-classification: In these cases, at least two features are crossed, and their commonalities and differences need to be identified (e.g. where does the banana fit best: to an apple, a pear, a ball of wool or a bucket?).

(4) Relationship identification: These tasks require that commonalities between relationships, for example, commonalities of sequences, are identified (e.g. arrange pictures of a comic strip story in the correct order).

(5) Relationship differentiation: In contrast to the previous category, differences between relationships need to be found out (e.g. disturbed sequence: objects are ordered according to increasing size except one object which does not follow this rule and which has to be sorted out).

(6) System composition: Both the commonalities and the differences of relations need to be identified (e.g. complete a matrix in which two features of objects need to be crossed).

Since the graphics of the paper-pencil-based training program I for preschool, primary school and special-needs students became out-dated displaying objects unknown to today's children (e.g. picture of a typewriter), Lenhard and colleagues (2012) developed a modern computer-based version. A further difference between the former paper-pencil-based version and the new computer-based version is that in the PC version, the training is embedded in a fantasy story: The learners are asked to help two elves to search for the "blue diamond of wisdom" (Lenhard et al., 2012).

In the following, two field studies are presented investigating the effectiveness of two forms of Klauer's cognitive training program I with special-needs students on cognitive and verbal development. In the first field study, a paper-pencil version of program I was applied in an Austrian sample, whereas in the second study, the computer-based version of program I was applied in a German sample. Within each study, it was assumed that:

1) the inductive reasoning training enhances fluid intelligence and results in increased intelligence test scores (*Hypothesis 1*);
2) due to the large amount of verbal activity, the training program should also result in improvements of verbal performance (*Hypothesis 2*);
3) the effects should be similar in an experimental group (first training group) and a waiting control group (second training group), that is, the increase in intelligence and verbal performance should be similar. This similarity of effects should ideally result in a 'rhomb pattern' of results (*Hypothesis 3*; cf. Figure 1). This is important in order show that the training results in an effect which is universal for all participants and rather independent of aptitude-treatment interactions (Cronbach & Snow, 1977);
4) the effects should be stable across a time period of six weeks (*Hypothesis 4*).

Comparing both studies, it was postulated that the computer-based version in which the cognitive training is embedded in a fantasy story (Lenhard et al., 2012) should be more effective than the paper-pencil version (*Hypothesis 5*) because the computer-based version should be more motivating and the objects should be easier to identify for the students.

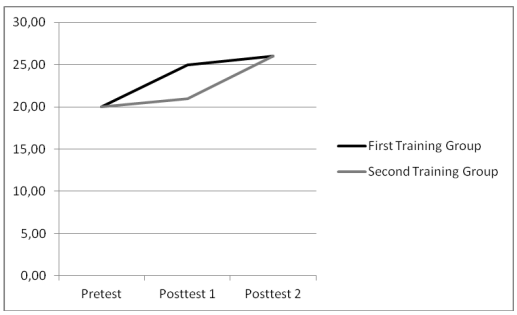


Figure 1. Hypothetical 'Rhomb Pattern' of Results of the Raw Scores.

Method

Across both studies, different trainers were used. However, within each study, the trainers also administered the tests.

Study 1

The study was conducted at three special-needs centers in Graz (Austria). The participants were nine girls and 25 boys at the age of 7;0 to 10;5 years (M = 8;9 years, SD = 1;1 years) visiting the first to fourth class of primary school. The parents and/or the teachers reported at least one of the following diagnoses (DSM-V) for each student, respectively: Autism spectrum disorder (e.g. Asperger syndrome); specific developmental disorder; expressive language disorder; reading disorder; disorder of written expression; mathematics disorder; conduct disorder; attention deficit hyperactivity disorder.

The children were matched to the first and second training group according to their pretest scores on the CFT 1 (Fiechtl, 2010). Cognitive ability was assessed in group settings by means of the three subtests Classifications, Similarities, and Matrices of the 'Culture Fair Test' CFT 1 (duration: 30-40 minutes; Cattell, Weiß, & Osterland, 1997), and verbal development was assessed by means of the four subtests Imitation of Grammatical Structure Forms, Generation of Morphemes, Sentence Generation, and Word Finding of the 'Heidelberger Sprachentwicklungstest' HSET (Grimm & Schöler, 1991), a German-speaking test which we adapted slightly to Austrian terms. Both tests were administered three times, as a pretest and as a first and second posttest. Due to several participants being ill in winter, less data were collected for the HSET. After the pretest, dyads of children of the first training group (n_1 = 17 children) took part in the cognitive training for children I by Klauer (1989). The training phase lasted for four weeks, with 2 sessions per week (each 30-40 minutes) and 15 tasks, respectively. Four weeks after the pretest, the first posttest was conducted. Only after this first posttest did the cognitive training start in the waiting control group (i.e., the second training group: n_2 = 17 children) for four weeks. After the waiting control group had finished the training, the second posttest was administered to all participants, twelve weeks after the pretest.

Results of Study 1

The effect size 'd_{corr}' in former evaluation studies on K. J. Klauer's cognitive training programs (e.g. Klauer & Phye, 2008) is corrected for pre-training differences (i.e., differences of the first post test are corrected for differences of the pretest, and differences of the second post test are corrected for differences of the first posttest; Cohen, 1988) and uses the pooled standard deviation. In addition, regression analyses are reported using the results of the previous test, respectively, and the experimental group (first training group vs. second training group) as predictors (see Table 1 for means and standard deviations).

On the first posttest, a small negative effect of the training on cognitive ability was found showing a superior performance of the waiting control group who did not train in this phase over the first training group (d_{corr} = -0.21, $\beta_{Train\text{-}Int}$ = -.17; see Figure 2). On the second posttest, the waiting control group who had received the training now, again scored slightly higher than the first training group who did not train in this phase, but the training effect also turned out to be slightly negative (d_{corr} = -0.13, $\beta_{Train\text{-}Int}$ = .02). The mean effect averaged across both training groups was d_{mean} = -0.17 ($\beta_{meanTrain\text{-}Int}$ = -.08).

A more inconsistent finding emerged for verbal development: On the first posttest, the waiting control group scored higher than the first training group (see Figure 3); thus, a small negative effect of the training on verbal ability was found (d_{corr} = -0.19, $\beta_{Train\text{-}Verb}$ = -.05); however, on the second posttest, the effect of the cognitive training on the second training group (i.e. the former waiting control group) was small but positive, (d_{corr} = 0.15, $\beta_{Train\text{-}Verb}$ = .13). This resulted in a mean effect of d_{mean} = -0.02, $\beta_{meanTrain\text{-}Verb}$ = .04).

Table 1: Means (and standard deviations) of the CFT 1 and the HSET scores used in Study 1.

	Pretest		Posttest 1		Posttest 2	
	CFT 1	HSET	CFT 1	HSET	CFT 1	HSET
First Training Group	22.41 (5.28) (n = 17)	31.24 (19.83) (n = 17)	24.59 (6.19) (n = 17)	40.56 (22.42) (n = 16)	25.76 (5.50) (n = 17)	39.47 (21.86) (n = 17)
Second Training Group	24.81 (7.38) (n = 16)	31.33 (20.60) (n = 15)	28.24 (4.68) (n = 17)	44.71 (23.59) (n = 14)	28.71 (5.30) (n = 17)	47.00 (19.61) (n = 17)

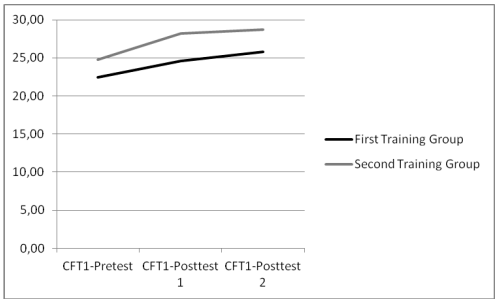


Figure 2. Raw Scores of the Intelligence Test Across Three Times of Measurement.

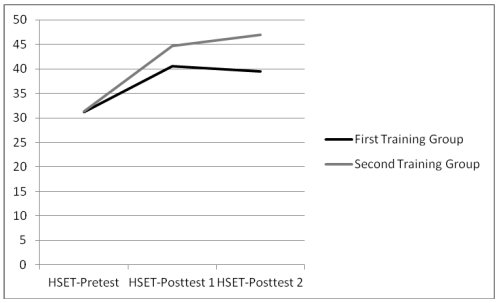


Figure 3. Raw Scores of the Verbal Competence Test Across Three Times of Measurement.

Thus, in contrast to Hypotheses 1 and 2, the paper-pencil-based version did neither support the inductive reasoning processes of special-needs students nor their verbal development. Consequently, no 'rhomb pattern' of results emerged, rejecting Hypothesis 3. Instead, the slight and relatively stable increases in cognitive and verbal ability (Hypothesis 4) can be traced back to (1) naturally occurring developmental effects and (2) a re-test or practice effect (i.e. getting used to work on the performance tests; Hasselhorn, 1995; Colom et al., 2010). In order to find out whether the modernized PC version (Lenhard et al., 2012) is more effective than the paper-pencil version (Hypothesis 5) of training program I (Klauer, 1989), Study 2 was also conducted with special-needs students.

Study 2

This study was conducted at a kindergarten with remedial education and at a special-needs school in Chemnitz (Germany). Seven girls and twenty boys participated who were about 5;2 to 9;7 years old (M = 7;8 years, SD = 1;5 years). As in Study 1, the parents and/or the teachers reported at least one of the following diagnoses (DSM-V) for each student, respectively: Autism spectrum disorder (e.g. Asperger syndrome); specific developmental disorder; expressive language disorder; reading disorder; disorder of written expression; mathematics disorder; conduct disorder; attention deficit hyperactivity disorder; learning difficulties.

The same tests as in Study 1 were used, that is, the CFT 1 measuring cognitive ability and the HSET assessing verbal development, as pretest and first and second posttest. The modernized, computer-based version of the cognitive training for children I (Lenhard et al., 2012) was conducted mostly in dyads (first training group: n_1 = 13; waiting control / second training group: n_2 = 14) with three units per week and ten tasks per unit (duration: 30 minutes, respectively). As in Study 1, the children were matched to the first and second training group according to their pretest scores on the CFT 1 (Jung, 2012; Voigt, 2012).

Results of Study 2

Again, the effect size 'd_{corr}' was used corrected for pre-training differences (i.e., differences of the first post test are corrected for differences of the pretest, and differences of the second post test are corrected for differences of the first posttest; Cohen, 1988), respectively, using the pooled standard deviation. Similarly, the regression analyses used the results of the previous test, respectively, and the experimental group (first training group vs. second training group) as predictors (see Table 2 for means and standard deviations).

On the first posttest, the first training group showed a superior performance than the waiting control group; thus, a small positive effect of the training on cognitive ability was found (d_{corr} = 0.20, $\beta_{Train\text{-}Int}$ = .10; see Figure 4). Similarly, in the second training group, a small positive effect of the training on cognitive ability was found on the second posttest (d_{corr} = 0.36, $\beta_{Train\text{-}Int}$ = .15). The mean effect, averaged across both training groups was d_{mean} =0.28 ($\beta_{meanTrain\text{-}Int}$ = .13) and equals to +4.2 IQ points.

In addition, a small positive effect of the training on verbal development was found in the first training group (d_{corr} = 0.12, $\beta_{Train\text{-}Verb}$ = .06), and also a small positive effect was found in the second training group (d_{corr} = 0.37, $\beta_{Train\text{-}Verb}$ = .18). This resulted in a mean effect of d_{mean} = 0.25, $\beta_{meanTrain\text{-}Verb}$ = .12).

Table 2: Means (and standard deviations) of the CFT 1 and the HSET scores used in Study 2.

	Pretest		Posttest 1		Posttest 2	
	CFT 1	HSET	CFT 1	HSET	CFT 1	HSET
First Training Group	19.46 (7.09) (n = 13)	38.87 (11.19) (n = 13)	22.38 (5.77) (n = 13)	41.26 (10.81) (n = 13)	21.85 (7.21) (n = 13)	42.28 (11.85) (n = 13)
Second Training Group	19.93 (7.07) (n = 14)	38.88 (9.10) (n = 14)	21.50 (7.35) (n = 14)	40.14 (8.06) (n = 14)	23.29 (5.31) (n = 14)	44.69 (7.52) (n = 14)

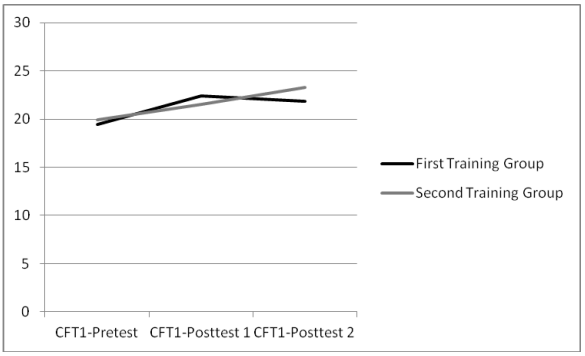


Figure 4. Raw Scores of the Intelligence Test Across Three Times of Measurement.

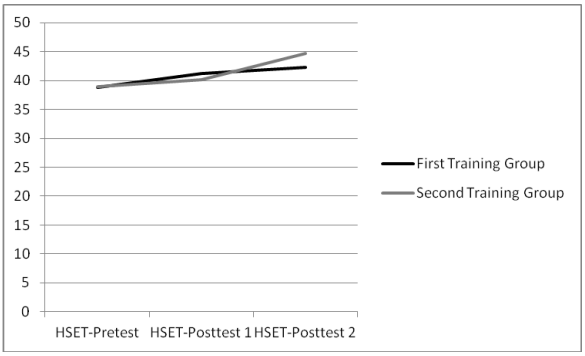


Figure 5. Raw Scores of the Verbal Competence Test Across Three Times of Measurement.

The PC version slightly improved the cognitive and verbal abilities of the special-needs students, supporting Hypotheses 1 and 2. Further, although the effects are not very pronounced, a slight 'rhomb pattern' of results was observable (see Figures 4 & 5); thus, the PC version should work relatively independently from further characteristics of the participants, supporting Hypothesis 3. The effects seem to be stable (Hypothesis 4); however, studies with more prolonged follow-up phases are needed in order to corroborate this assumption. Finally, comparing the paper-pencil-based (Study 1) and the PC version (Study 2) of the training, the latter version was more effective, confirming Hypothesis 5.

Discussion

Inductive reasoning is the ability to identify similarities and differences of attributes and relations (Klauer & Phye, 2008), and it is a core feature of human intelligence (Klauer & Phye, 2008; Lenhard & Lenhard, 2011; Rindermann & Baumeister, 2013). Inductive reasoning trainings for all age groups have been developed and tested by K. J. Klauer. His training programs can be characterized as a German approach to cognitive apprenticeship (e.g. Collins et al., 1989) because the classical methods of modeling, coaching, scaffolding, fading, articulation, reflection, and exploration are applied.

In two field studies, we tested the effectiveness of the paper-pencil-based version of program I and compared it to the effectiveness of the PC version with regard to improving the cognitive and verbal abilities of special-needs students. The first field study (N_1 = 34; paper-pencil version) was conducted in Austria, the second field study (N_2 = 27; PC version) in Germany.

In the first field study, no positive effect of the training on intelligence and verbal development was found. Instead, all students improved their cognitive and verbal abilities across the three times of measurement independently of the cognitive training. One reason why no positive training effect was found could be that the tasks were not motivating enough for the students. In addition, the training tasks were similar to the tasks of the intelligence test (CFT 1) which resulted in boredom and decreased motivation during the test phase. That is, although the training and performance test tasks seemed to be too easy for the students, no ceiling effect of their cognitive and verbal performance was observable. Therefore, further studies should use broader intelligence tests (e.g. the Wechsler Intelligence Scale for Children; Wechsler, 2004) which offer farther transfer between training tasks and test tasks.

In the second field study, a PC version of the training program I was used (Lenhard et al., 2012). In this PC version, out-dated graphics of the paper-pencil version had been exchanged, and a motivating fantasy story had been built around the cognitive training tasks. Thus, this PC version is even more typical of a cognitive apprenticeship approach than the previous version because situated learning is more strongly applied when the students participate in the fantasy story. The PC version had a small and positive impact on the students' cognitive and verbal abilities, over and above the developmental gains which were also observable for all participating students. Although only a small effect of the cognitive training was found on the students' intelligence, this small effect is still practically meaningful because an increase of four IQ points corresponds to more than half a year of schooling and maturation combined (Rindermann, 2011)!

Since our small training effects are in contrast to those reported by Klauer and Phye (2008), it is important to identify moderating variables (e.g. type of developmental disorder, IQ levels, motivation etc.) which support or hinder training gains. In addition, it is important to include studies like our two in future meta-analyses in order to reduce publication biases. Different assumptions exist regarding the question for which target groups such cognitive trainings would be especially effective: According to Klauer and Phye (2008), students with learning disabilities should profit more than students without learning disabilities (w.r.t. intelligence: d = 0.54 vs. d = 0.52; w.r.t. academic performance: d = 0.94 vs. d = 0.69; Klauer & Phye, 2008). Similarly, large effects of cognitive trainings were shown in developmental countries (e.g. Irwing et al., 2008: Sudan, d = 0.47, this equals to 7 IQ points). In further own studies, however, we found larger effects of Klauer's training for students who show higher academic performance and for elderly people with higher levels of education compared to less educated seniors of the same age (Rindermann & Baumeister, 2013) – thus, 'Matthew effects' ("the rich get richer, and the poor get poorer"; Matthew 25:29) seem to be more common for the cognitive training programs by Klauer than ceiling effects.

Limitations of both studies were the small sample sizes. Further, there was no follow-up phase for the second training group. Therefore, it is possible that in the first study, positive effects of the training would have emerged in later phases after the students had more opportunities to apply the trained skills in daily life. Future studies should try to investigate the training's effectiveness with larger samples and longer follow-up phases (e.g. 3 to 15 months; Klauer & Phye, 2008).

Finally, a problem of the statistical method of effect estimation is that different effect sizes result in different interpretations: Several variants for calculating d-values exist depending on the homogeneity or heterogeneity of the sample and the resulting standard deviations (cf. Rindermann & Baumeister, 2013). In addition, d-values as effect levels are numerically higher than r- or β-values (e.g. small effect according to

Cohen, 1988: d = .20 vs. r / β = .10). Therefore, it is recommended to report at least two variants for calculating the effect size, for example, d and β. If the performance test (e.g. intelligence test) was conducted completely, test norms should be used to quantify the effectiveness in IQ points based on the population variance.

A further methodological challenge is to use a suitable control group, for example, in the form of a further training program pursuing different aims (e.g. Jaeggi et al., 2011) or by separating and comparing the different components of one and the same training program (e.g. Dorbath, Hasselhorn, & Titz, 2011).

Since K. J. Klauer's cognitive training programs are the most often evaluated ones with more than 100 experimental studies and more than 4.000 participating children from Europe and the USA, the potential effects are undoubted (Klauer, 2014). Thus, it can be recommended to integrate the training in the curriculum of kindergartens and schools, including refreshing sessions several months later (Möller & Appelt, 2001) – this would take no longer than 10 lessons and could substantially help children to improve their cognitive abilities and academic performance.

References

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Butera, F., Caverni, J. P., & Rossi, S. (2005). Interaction with a high- versus low-competence influence source in inductive reasoning. *The Journal of Social Psychology, 145,* 173-190.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65*, 684-698.

Cattell, R. B., Weiß, R. H., & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)*. Braunschweig: Westermann.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*, 703-722.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Colom, R., Quiroga, M. A., Shih, P. C., Martínez, K., Burgaleta, M., Martínez-Molina, A., et al. (2010) Improvement in working memory is not related to increased intelligence scores. *Intelligence, 38,* 497-505.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: a handbook for research on interactions.* New York: Irvington.

Dorbath, L., Hasselhorn, M., & Titz, C. (2011). Aging and executive functioning: A training study on focus-switching. *Frontiers in Psychology, 2*, 1-12.

Fiechtl, S. (2010). *Effekte kognitiven Trainings bei Kindern mit Defiziten in deren schulischen Fertigkeiten* (Unpublished diploma thesis). University of Graz, Graz.

Grimm, H. & Schöler, H. (1991). *HSET – Heidelberger Sprachentwicklungstest*. Göttingen: Hogrefe.

Hasselhorn, M. (1995). Kognitive Trainings: Grundlagen, Begrifflichkeiten und Desiderate [Cognitive training: Basics, concepts, and desiderata]. In W. Hager (Ed.), *Programme zur Förderung des Denkens bei Kindern* [Programs for promoting thinking in children] (pp. 14-40). Göttingen, Germany: Hogrefe.

Hendricks, C. (2001). Teaching causal reasoning through cognitive apprenticeship: What are the results from situated learning? *The Journal of Educational Research, 94,* 302-311.

Irwing, P., Hamza, A., Khaleefa, O., & Lynn, R. (2008). Effects of Abacus training on the intelligence of Sudanese children. *Personality and Individual Differences, 45*, 694-696.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 10081-10086.

Jung, C. (2012). *Effekte des computergestützten Denktrainings nach Klauer auf die kognitiven Fähigkeiten von Kindern* (Unpublished bachelor's thesis). Technische Universität Chemnitz, Chemnitz.

Klauer, K. J. (1989). *Denktraining für Kinder I. Ein Programm zur intellektuellen Förderung.* Göttingen: Hogrefe.

Klauer, K. J. (1991). *Denktraining für Kinder II.* Göttingen: Hogrefe.

Klauer, K. J. (2008). *Denksport für Ältere: Geistig fit bleiben* (2., überarbeitete Auflage). Bern: Verlag Hans Huber.

Klauer, K. J. (2014). Training des induktiven Denkens – Fortschreibung der Metaanalyse von 2008 [Training of inductive reasoning – an updated meta-analysis]. *Zeitschrift für Pädagogische Psychologie, 28,* 5-19.

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research, 78*, 85-123.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71,* 449-521.

Lenhard, A., & Lenhard, W. (2011). Computerbasierte Intelligenzförderung mit den „Denkspielen mit Elfe und Mathis" – Vorstellung und Evaluation eines Computerprogramms für Vor- und Grundschüler. *Empirische Sonderpädagogik, 2,* 105-120.

Lenhard, A., Lenhard, W., & Klauer, K. J. (2012). *Denkspiele mit Elfe und Mathis. Förderung des logischen Denkvermögens für das Vor- und Grundschulalter* [*Thinking games with Elfe and Mathis. A training of logical thinking skills for preschool and primary school age*]. Göttingen: Hogrefe Verlag.

Marx, E. (2005). Kognitive Entwicklungsförderung bei hörgeschädigten Kindern [Cognitive competence development in hearing impaired children]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 37*, 36-45.

Marx, E. (2006). Kognitive Förderung Jugendlicher mit Lernstörungen. *Psychologie in Erziehung und Unterricht, 53*, 166-177.

Masendorf, F., & Klauer, K. J. (1987). Intelligenztraining bei lernbehinderten Sonderschülern: ein kognitiv-prozessualer Ansatz [Cognitive training with special-needs students with learning disorders]. *Psychologie in Erziehung und Unterricht, 34,* 14-19.

Möller, J., & Appelt, R. (2001). Auffrischungssitzungen zur Steigerung der Effektivität des Denktrainings für Kinder I [Booster sessions to foster the effectiveness of the training to reason for children I]. *Zeitschrift für Pädagogische Psychologie, 15,* 199-206.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review, 6*, 321-349.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction, 1*, 117-175.

Rindermann, H. (2011). Intelligenzwachstum in Kindheit und Jugend [Increase of intelligence in childhood and youth]. *Psychologie in Erziehung und Unterricht*, 58, 210-224.

Rindermann, H., & Baumeister, A. E. E. (2013). Effekte des Klauerschen Denktrainings bei Schülern mit Migrationshintergrund, Förderbedarf und bei Senioren [Effectiveness of Klauer's cognitive training with immigrant students, special-needs students and elderly people]. *Psychologie in Erziehung und Unterricht*, 60, 121-132.

Rindermann, H., Flores-Mendoza, C., & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences, 20*, 544-548.

Rindermann, H., Michou, Ch. D., & Thompson, J. (2011). Children's writing ability: Effects of parent's education, mental speed and intelligence. *Learning and Individual Differences, 21*, 562-568.

Rogoff, B., Paradise, R., Mejía Arauz, R., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology*, *54*, 175-203.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience, 2,* 1-10.

Schoenfeld, A. H. (1985). *Mathematical problem solving.* New York: Academic Press.

Sonntag, W. (2004). Experimentelle Untersuchungen zum Einfluss des Klauerschen Denktrainings auf mathematisches Denken und Lernen von lernbehinderten Sonderschülern [Experimental studies on the impact of Klauer's training to reason on mathematical thinking and learning with slightly retarded students]. *Zeitschrift für Pädagogische Psychologie*, *18*, 101-111.

Sonntag, W. (2006). Auffrischungssitzungen zur Steigerung der Effektivität des "Denktrainings für Jugendliche" bei lernbehinderten Sonderschülern [Do booster sessions increase the effects of an inductive training on intelligence and learning with slightly retarded children in special schools?]. *Psychologie in Erziehung und Unterricht, 53*, 178-187.

Voigt, I. (2012). *Evaluation der Effekte des computergestützten Denktrainings nach Klauer auf die Sprachfähigkeit von Kindern* (Unpublished bachelor's thesis). Technische Universität Chemnitz, Chemnitz.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Wechsler, D. (2004). *The Wechsler intelligence scale for children - fourth edition*. London: Pearson Assessment.

Acknowledgments

We want to express our gratitude to Sylvia Fiechtl, Cindy Jung, and Isabel Voigt for collecting the data.

Development of an Empirically-Based Learning Performances Framework for 3rd-Grade Students' Model-Based Explanations about Hydrologic Cycling

Cory T. Forbes, University of Nebraska-Lincoln, Lincoln, NE 68583, cforbes3@unl.edu
Christina Schwarz, Michigan State University, East Lansing, MI 48824, cschwarz@msu.edu
Laura Zangori, University of Nebraska-Lincoln, Lincoln, NE 68583, laura.zangori@huskers.unl.edu

Abstract: Elementary students should engage in the articulation, negotiation, and revision of model-based explanations. However, scientific modeling remains underemphasized in elementary science learning environments and more research is needed to understand early learners' engagement in domain-specific modeling practices. To address this need, we are engaged in design-based research to foster and investigate 3rd-grade students' model-based explanations for hydrologic phenomena. First, we developed an empirically-based learning performances framework that integrates relevant science content and modeling practices. This framework a) grounds the iterative adaptation and enhancement of a commonly-used curricular unit and b) lays the foundation for ongoing development of an associated learning progression. Second, we report on findings from analysis of 3rd-grade students' model-based explanations around the water cycle. Results indicate that elementary students generate mechanism-based causal claims and highlight target concepts and modeling practices emphasized in students' model-based explanations for hydrologic cycling.

Study Rationale and Contribution

To become scientifically-literate, students must learn to reason about complex, global issues such as water resource management and sustainability at an early age (American Association for the Advancement of Science [AAAS], 2007; National Research Council [NRC], 2012; 2007). Early learners' understanding of the nature of water, how it cycles and changes state, and its relationship to human activities, are all necessary to help them make sense of everyday experiences and to serve as a foundation for their learning about other Earth systems and water-related global issues with scientific, social, and economic dimensions. However, past research has shown that early learners often struggle to understand hydrologic phenomena (e.g., Bar, 1989; Henriques, 2002). Elementary students therefore need greater support for learning about hydrologic systems.

To develop conceptual understanding of hydrologic systems, students must engage in theory-driven scientific practices focused on the articulation, negotiation, and revision of model-based explanations (Braaten & Windschitl, 2011; NRC, 2007; Windschitl, Thompson, & Braaten, 2008). Modeling is a core scientific practice advocated in the *Next Generation Science Standards* (NRC, 2012) in which models, or working representations of complex natural systems, are used to reason scientifically about system-specific phenomena. Modeling practices, however, remain underemphasized in K-12 science, particularly in the elementary grades, despite growing evidence that, with scaffolding, elementary students can effectively engage in scientific practices (Hapgood, Magnusson, & Palinscar, 2004; Hardy, Jonen, Möller, & Stern, 2006; Herrenkohl & Cornelius, 2013; Lehrer & Schauble, 2006; Manz, 2012; Metz, 2004; McNeill, 2011). Investigations of elementary students' model-based reasoning, particularly about the water cycle, are largely absent from the literature. More research is therefore needed to inform the design of elementary science learning environments that afford students opportunities to develop and use models to formulate explanations.

To begin to address these issues in the field, we are engaged in three years of exploratory research and development to foster and investigate 3rd-grade students' model-based explanations about the water cycle. We draw upon work on scientific modeling (Lehrer & Schauble, 2006; Schwarz et al., 2009; Windschitl, Thompson, & Braaten, 2008), content- and practice-based learning progressions (Alonzo & Steedle, 2008; Lee & Liu, 2010; Mohan, Chen, & Anderson, 2009; Stevens, Delgado, & Krajcik, 2010), and heuristics for curriculum materials development (Krajcik, McNeill, & Reiser, 2007; Shin, Stevens, & Krajcik, 2010) to articulate a domain-specific student learning performance framework that integrates science content and scientific practice (i.e., modeling). Such learning performances are critical to ground the design of curricular, instructional, and assessment dimensions of classroom interventions. In this paper, we report on work from Year 1 of the project: a) the empirical development of learning performances and b) analysis of student artifacts to investigate their use of models to formulate evidence-based explanations for hydrologic cycling. We ask two research questions: 1) *what are measureable levels of 3rd-grade students' model-based explanations about water?* and 2) *how do 3rd-grade students formulate model-based explanations for target concepts related to water?* This work foregrounds elementary students' learning as constituent component of their discipline-specific epistemic practices and therefore exemplifies the ICLS 2014 conference theme of 'learning and becoming in practice'.

 © ISLS

Theoretical Framework

Students' Model-Based Reasoning

The hydrologic cycle is a foundational model-based scientific concept highlighted throughout the K-12 science curriculum (AAAS, 2007; NRC, 2012). Models are defined as abstracted, multi-modal representations of natural systems, not exact recreations, which are used within communities to illustrate, predict, and explain system-specific scientific phenomena. They are used extensively by hydrologists, climate scientists, meteorologists, and soil scientists to make predictions about, investigate, and explain hydrologic cycling. Past research has shown, however, that students possess a diverse set of pre-existing ideas about hydrologic phenomena, ideas that are often times inconsistent with scientific explanations (e.g., Bar, 1989; Henriques, 2002). Model-based investigative practices can support students in constructing, negotiating, and revising explanations for scientific phenomena in a variety of scientific disciplines. Students' *construction*, *evaluation,* and *revision* of models of hydrologic cycling can help them make their thinking visible, but such models also serve to shape their reasoning about water systems through *use*. As such, models act as both *representations* and *tools*, not only serving as records or artifacts of sense-making activity, but also playing a critical role in shaping reasoning activity itself. Past research has shown that elementary students often have difficulty engaging in model-based reasoning around scientific phenomena (Lehrer & Schauble, 2006; Schwarz et al., 2009). They may emphasize singular events or phenomena in their models rather than interacting systems. Even when students do focus on broader systems, they may not connect system-specific phenomena to empirical data. While developmental limitations are often viewed as obstacles to young students' learning, the authors of *Taking Science to School* (NRC, 2007) note that "young children have a repertoire of cognitive capacities directly related to many aspects of scientific practice, and it is problematic to view these simply as a product of…development" (pg. 44). Recent empirical research provides evidence that, with scaffolding, early learners can engage productively in scientific modeling (Lehrer & Schauble, 2006; Manz, 2012).

An Integrated Learning Performances Framework for Model-Based Explanations

A comprehensive framework is required to both foster and assess students' formulation of model-based explanations in effectively-designed elementary science learning environments. Empirically-tested learning progressions have been developed to account for students' conceptual understanding in various content domains (e.g., Alonzo & Steedle, 2008; Lee & Liu, 2010; Mohan, Chen, & Anderson, 2009; Stevens, Delgado, & Krajcik, 2010). However, knowing and doing are mutually constitutive – what the learner knows influences what he/she does, and vice versa. The practice-based nature of learning is encapsulated by *learning performances* (Krajcik, McNeill, & Reiser, 2007; Shin et al., 2010), or behavioral claims that specify how students engage in scientific practices to employ their conceptual knowledge. A learning progression is comprised of individual learning performances that represent domain-specific scientific practices. The learning progressions community has begun to acknowledge the need for learning progressions that not only account for students' *conceptual understanding*, but also their engagement in *scientific practices* (e.g., Schwarz et al., 2009; Shin et al., 2010). While a small number of researchers have explored practice-based learning progressions, including those for elementary students' modeling practice and model-based reasoning (Lehrer & Schauble, 2006; Schwarz et al., 2009), much more work is needed to articulate practice-based learning progressions that account for BOTH epistemic and conceptual dimensions of elementary students' domain-specific learning.

We have generated a hypothetical learning performances framework for students' use of scientific models to formulate explanations for target concepts around the big idea that all geosystems are the result of energy flow and mass cycling (AAAS, 2007; NRC, 2012). As related to water, the three concepts underlying the big idea targeted in this project are *water exists in different forms below, at, and above the Earth's surface* (Concept 1); *water on Earth is in motion and cycles at a global scale* (Concept 2); and *the cyclical movement of water on Earth shapes and impacts the geosphere* (Concept 3). Students are afforded opportunities to generate and use models of the water cycle to formulate evidence-based explanations for each of these concepts. Epistemic dimensions of scientific explanation include *components*, *sequence*, *explanatory process*, *mapping*, and *principle* (Schwarz et al., 2009). The learning performances framework is shown in Table 1.

Table 1: Learning performances framework for students' model-based explanations about hydrologic cycling

	(1) Forms of water	(2) Water in motion	(3) Water/geosphere interactions
Components	The properties of water enable it to exist in 3 forms - liquid, vapor/gas, and ice – depending on temperature. Temperature affects the state of water. This relationship explains why	The cyclic motion of water above, on, and within the Earth is largely determined by the force of gravity and geospheric components	Just as water's movement is influenced by the geosphere, water in turn shapes the geosphere as it moves over and through it. Many landforms and geospheric features we observe everyday are
Sequence			
Explanatory Process			
Mapping			
Principle			

Components	water is found all over the Earth in different forms.	with which water interacts.	a result of these processes.
Sequence			
Explanatory Process			
Mapping			
Principle			

Together, these features comprise mechanism-based explanations (Braaten & Windschitl, 2011; NRC, 2012) for water-related phenomena. The framework in Table 1 foregrounds how students attend to scientific 'mechanisms', or the "unobservable, theoretical components" (Braaten & Windschitl, 2011, p. 662) that bring about an observable effect. The *components* feature emphasizes both visible and non-visible elements of the phenomena. *Sequences* establish temporal relations between system sub-processes. *Explanatory process* emphasizes mechanisms that explain process sequences. *Principle* involves a generalization about the phenomena that relates to abstracted components of the model. Finally, *mapping* emphasizes explicit statements that explain how the representation or components in the representation relates to the physical phenomenon. A model-based explanation occurs when students use systems representations to build on their existing knowledge to understand 'how' and 'why' they observed what they did (e.g. the mechanism). Consistent with this view, the purpose of explanation construction in the science classroom is for students to make sense of how the world works by connecting the cause and effect of natural phenomena with its underlying mechanism (explanation).

Method

The 3-year project is grounded in design-based empirical research. Empirical findings are used in an iterative manner to inform the design and implementation of curricular and instruction interventions to foster students' construction of model-based explanations about water.

Participants and Context

In Year 1 of the project, six 3rd-grade elementary teachers were recruited from schools in a Midwestern state. 179 eligible 3rd-grade teachers taught kit-based elementary science curriculum materials provided and managed by a regional educational services unit (ESU). Project teachers were selected using purposeful, maximum-variation sampling (Patton, 2001) in consultation with ESU staff to identify experienced teachers from urban, rural, and suburban settings serving students from underrepresented demographic groups with widely variant socio-economic profiles. Participant teachers already used the Full Option Science System (FOSS) Water module through the ESU. While FOSS units are reasonably well-developed, the Water module does not engage students in substantial modeling to situate domain-specific conceptual understanding targeted in unit investigations within broader, systems understandings of hydrologic cycling. We therefore engage in *construct-centered design* (CCD - Shin et al., 2010) and use the learning performances framework to enhance the FOSS Water module to more effectively foster students' model-based explanation-construction. CCD is based upon empirically-tested heuristics for learning progressions-based curriculum development (Krajcik, McNeill, & Reiser, 2007) and involves four steps. First and second, we define the content and articulate learning performances (Table 1) that provide the conceptual and epistemic underpinnings of module development. Third, we use learning performances developed in Step 2 to design an accompanying modeling task that is integrated into the existing module. In Year 1, we employed a limited version of the modeling task where students construct models at the beginning at the end of the unit only. Fourth, the full version of the modeling task will be developed and implemented in Years 2 and 3 based upon empirical results from Year 1.

Data Collection

Data collection occurred during the enactments of the modified FOSS Water module in each participant teacher's classroom. To investigate 3rd-grade students' model-based explanations for water, we draw upon a number of data sources. First, we collected pre- and post-unit student modeling tasks (n_{pre}=112, n_{post}=107). The modeling task is designed to elicit student learning performances in Table 1 through diagrammatic, concept-process models in which students draw upon a variety of text-, numeric-, and image-based elements. Students are prompted with the question, 'What happens to rain when it reaches the Earth's surface?' and asked to use their existing ideas to *construct* a model of groundwater cycling. Students concluded the module by *evaluating* and *revising* their models. Students were also asked to justify their modeling decisions through scaffolded written responses. Second, clinical interviews were conducted with five students from each classroom in conjunction with their pre- (n=30) and post-unit models (n=30). Students were purposefully selected (Patton, 2001) by the research team and the six participant teachers to represent a continuum of academic achievement and classroom engagement. The clinical interview protocols were designed to elicit student reflections on their water cycle models and written responses around each of the learning performances in Table 1.

Data Analysis

Data analysis involved both quantitative and qualitative methods. All audio recorded interview data was transcribed, student artifacts scanned and digitized, and all data imported into ATLAS.ti, a widely-available qualitative data analysis suite, for coding. Student interview data was coded using the learning performances framework in Table 1. This coding process involved 15 codes, one for *components*, *sequence*, *explanatory process*, *principle*, and *mapping* for each of the three target concepts. Joint coding was performed on a 10% data sample. Inter-coder reliability was 85% before discussion and 100% after discussion. Code queries were conducted to isolate data for each of the codes representing learning performances in in Table 1. Within each data subset, analyses focused on the identification and articulation of measureable 'levels' for each component of students' model-based explanations about the three concepts in Table 1. Qualitative analysis involved an iterative process of data reduction, displaying, and verification to identify learning performance levels. The levels of a learning performance represent varying degrees of sophistication for domain-specific, model-based explanation-construction. The learning performance levels were then used to develop a scoring rubric as a scaled measure of students' understanding of the three target concepts as evidenced in the models and written responses in the modeling task. The rubric allowed for examination of the learning performances at three levels of sophistication (Levels: 0 – 3). Students' pre-unit and postunit modeling tasks were scored using the rubric. The student data was nested per teacher which required a multi-level model analysis (Littell, Milliken, Stroup, Wollonger & Schabenger, 2006). The analysis was conducted in SAS using a double-factor repeated-measures mixed-model ANOVA. The dependent variable was the postunit models while the independent variables were the dimension, concept, and the interaction between dimension and concept. The ANOVA formula is $Yijk = \mu + \alpha j + \beta k + (\alpha\beta)jk + eijk$ where i is the pre or postunit models, j is concept, k is dimension, and e is the error in the dependent and independent variables. Statistical-significance of the interaction term indicates that some combination of dimension and the concepts affect the models (Kleinbaum et al., 1998).

Results

Defining Levels of Learning Performances

Analysis of the student interviews yielded three definitive levels for each of the learning performances illustrated in Table 1. In this paper, we discuss learning performance levels for each of the epistemic features of model-based explanation in Table 1 for Concept 1 to illustrate ranges of students' model-based explanation-construction and distinguishing characteristics of the learning performance levels.

For a model to illustrate dynamic interactions within a system, it must first include the relevant *components*. For Concept 1, students should recognize that water exists in three forms – liquid, ice, and vapor – and that water is represented in these three forms at various points throughout the water cycle. Further, an important underlying characteristic of mechanism-based explanations is accounting for unseen/unobservable forces that drive observable cause and effect. The three levels of the *components* learning performance for Concept 1 emphasize student understanding of both visible and non-visible phases of water. At Level 1, students include at least one representation of VISIBLE water in a naturally occurring form (rain, surface water, ice, clouds, etc.). At Level 2, students include multiple representations of both visible and non-visible water. Non-visible water forms would include water vapor. Finally, at Level 3, students include multiple visible and non-visible water forms, including subsurface groundwater. Level 3 understanding is illustrated by the following interview excerpt:

> S: When the water hits a tree it grows and then when it takes it out, it takes water from the tree and the surface runoff I did a tiny mountain right there and then the water falling down.
> I: That's your hill with the surface runoff?
> S: Then the rain goes up to the ground and then it goes up like the last one. Its evaporation, condensation and then precipitation happens when there's cold and hot air up there. When it starts to get plus and minuses there's lightning and thunderstorm and everything. (P:46:31:42)

At a Level 3, students identify both visible and non-visible atmospheric and geospheric components of the water cycle which, as illustrated in the sample student quote, is an enabling factor in their articulation of process and mechanism for hydrologic system dynamics.

For a model to illustrate mechanisms for system processes, it must first also include the relevant *sequence* of these processes and their constituent parts. For Concept 1, students should recognize that water goes through phase changes between its three forms – liquid, ice, and vapor – and that phase change occurs throughout the water cycle. The three levels of the *sequence* learning performance for Concept 1 emphasize student understanding of both what phase changes occur and the directionality or order in which they occur. For the *sequence* dimension of Concept 1, Level 1 understanding involves simple description of at least one phase

change, for example, from water vapor to liquid water in the form of precipitation. At Level 2, students began to recognize multiple phase changes that occur at places in the water cycle and how one sequentially leads to the next. However, students do not illustrate understanding that phase changes can occur in multiple directions. At a Level 3, students begin to describe phase change in multiple directions. Level 3 understanding is illustrated by the following interview excerpt:

> I: … do you think there is water up in the sky all of the time? How does it get up there?
> S: Yes, it evaporates...which means that little tiny, that you can't see, water droplets come out of the oceans, lakes, rivers, and ponds, and they go up into the clouds. And, when there is too many water droplets in the clouds, it starts to rain or snow. (P6:48:55)

In the Level 3 example student quote, the student describes both evaporation and precipitation simultaneously. While condensation is also an important process in part of the water cycle, no student described it as a stand-alone, mediating phenomena that leads to cloud formation.

Building upon students' model-based reasoning of both system *components* and *sequence*, students should articulate a mechanism-based *explanatory process* for system processes represented in their models. For Concept 1, students should recognize that varying temperature impacts phase change as an unseen causal mechanism. The three levels of the Concept 1 *explanatory process* learning performance emphasize student understanding of relationships between related constructs and their underlying mechanisms. For the *explanatory process* dimension of Concept 1, Level 1 understanding involves simple description of the process, such as components or sequences, indicative of the absence of a formal mechanism, without which there is no explanation-construction occurring. For example, students might recognize ice in polar regions and that it can melt without referencing temperature. At Level 2, students began to recognize a relationship between temperature and phase change. At this level, it is an associative relationship, however, not a causal one, in that students may associate changing temperature with phase change but not attribute one to the other. At a Level 3, students begin to represent temperature as a direct, causal agent that explains observable phase change as part of the water cycle. Students' explanations are grounded in models that represent temperature in some way as a critical element of the water cycle. Level 3 understanding is illustrated by the following interview excerpt:

> S: I put the pluses for hot and I put some pluses there you can't see the minuses much that I put up there for cold. That also creates rain, it helps rain fall.
> I: The minuses, the cold helps rain fall?
> S: Yes but usually the hot on the bottom. Yes but when the hot is on the top of the cold on the bottom that actually creates rain too. See? Usually it's the other way around and that's also why it creates rain. (P 1: 3053:3059)

As illustrated in the excerpt, the student describes heat causing evaporation and cooling causing condensation/precipitation, which he/she represented in the model as +s and –s.

A critical element of model-based explanations involves students' making connections between their representation of the phenomena and the phenomena itself. This component, referred to as *mapping*, involves students explicitly articulating how their model, or elements of their model, relates to the natural system. The three levels of the mapping learning performance for Concept 1 emphasize students' relating their representations of forms of water to the natural world. At Level 1 understanding, students simply state that some aspect of their model is intended to represent some component of the system. At Level 2, students provide some rationale for how or why their model represents the natural system. For example, students may state that clouds are grey because they're full of water as they have observed on rainy days. At a Level 3, students articulate explicit rationales for representational elements of their models and how they map onto the real-world phenomena. Level 3 understanding is illustrated by the following interview excerpt:

> I: You said condensation was something you learned about in class…where is that in your model?
> S: In the test we did in [class], I drew the cloud and cut the cloud in half and showed you the little water droplets in it. That side had more. That side had less. Then I made a little circle and a little line to connect it with a bigger circle to show that when it was condensing and …two of the little gas water droplets were coming together and getting bigger in there and turning into rain drops and falling. (P 4: 1574:1599)

In the Level 3 example student quote, the student describes how he/she represented evaporation and condensation to illustrate phase change underlying these processes.

Finally, an important aspect of model-based explanations involves students generalizing from their model about an underlying scientific *principle*. While specific *components* and *sequences* represented in the model *map* onto real-world phenomena, and model elements illustrate *explanatory processes* for these systems, students should also be able to derive generalized principles about system-specific phenomena. The three levels

of the *principle* learning performance for Concept 1 emphasize students' relating their representations of forms of water to forms of water in the natural world. At Level 1 understanding, students may identify a generalized scientific principle, but do so either erroneously or inconsistently. For example, students might attribute gravity to phase change or incorrectly state that increasing temperature causes water to freeze. Level 2 understanding involves some accurate articulation of a relevant scientific principle but relating it incompletely. At Level 3, students correctly and fully match components of the model with an underlying scientific principle. Level 3 understanding is illustrated by the following interview excerpt:

> S: Usually, water that stays there usually goes deeper into the ground.
> I: Why does it go deeper?
> S: The big drops probably are heavier so it goes down faster so it, so like the air can't push it up again. And the small drops go down slower because and because they're lighter and they're, um, the air can push them up easier (P24: 827:881)

In the Level 3 example quote, the student describes a fundamental cause and effect relationship that explains the presence of subsurface groundwater – that the downward force of water in large volumes can more easily move through Earth materials. However, the student does not attribute this to gravity or other explanatory processes, thus illustrating principle as an epistemic commitment distinct from explanatory process.

Students' Model-Based Explanations for Hydrologic Cycling

To address research question #2, we analyzed students' pre- and postunit modeling tasks for the five epistemic features of mechanism-based explanation construction (Table 1). First, analysis suggests that there was no statistically-significant difference between epistemic features (Table 1) represented in the students' pre- and postunit models. However, a statistically-significant difference was observed between pre- and postunit models for each Concept 1, $F(2,3147) = 7.12$, $p < .0001$, Concept 2, $F(2,3147) = 6.77$, $p < .0001$, and Concept 3, $F(2,3147) = 21.61$, $p < .0001$. These results indicate that while the students did not include additional epistemic features within each concept from pre- to postunit models, the features they did represent for each concept increased in sophistication over the course of the unit. In the postunit models, we found a significant interaction between features of epistemic dimensions and concepts, $F(8, 1355) = 45.67$, $p = <0.0001$, when controlling for the pre-unit model scores. We used paired-samples t-tests to make *post hoc* comparisons between concepts for each feature to ascertain which epistemic dimension contributed most to differences observed between target concepts in postunit modeling tasks. The results indicate statistical significance for *mapping*, $t(1355) = -5.16$, $p = <0.0001$, and *sequence*, $t(1355) = -11.53$, $p = <0.0001$, between Concept 1 and 2. Students' engaged in more sophisticated representations of Concept 2 ($\bar{x} = 1.08$) than Concept 1 ($\bar{x}=0.76$) due to a stronger emphasis on representation of *sequences* and *mapping* for Concept 2. Further, differences between epistemic features for Concept 3 compared to both Concept 1 and 2 were all statistically-significant. This result is due to the less frequent representation of Concept 3 ($\bar{x} =0.08$) in students' models. Together, these results suggest that students' models emphasized water movement more so than the forms of water. Further, representations of interactions between water and Earth materials (Concept 3) were rare. Students' representations of target concepts increased in sophistication during the unit. Differences in representations of target concepts in students' postunit models were largely attributable to two epistemic features: *sequence* and *mapping*.

Findings from qualitative analysis for research question 2 illustrate two dominant themes. First, findings show that students' efforts at *mapping* most frequently emphasized water in motion above ground and their connection of their representations to the physical world did not typically change from pre to postunit models. Students expressed that their 2-D drawings were insufficient for 'realistic' representations of precipitation and were concerned that they were unable to represent all of the pieces of this process that they felt were necessary in the 'real world'. This was most frequently observed in students' attempts to represent the presence of the sun prior to rain, in which students would draw the sun, and then the 'absence' of the sun during rain, which often involved students drawing over the sun. When asked why the sun was initially present and then drawn over they identified that in the 'real' world, the sun is originally visible but "when it rains, it gets all foggy and clouds start to turn gray and then it [the clouds] just covers up the sun" (P 1: 3377:3386). In their manner of including then removing the sun, they were attempting to map accurately the processes they observed occur during precipitation rather than identifying their drawings as abstractions of the physical world (Figure 1).

Second, we found that *sequences* underlying students' model-based explanations relied on combining scientifically-accepted mechanisms (e.g. gravity, temperature) with their pre-existing mechanisms (e.g. alternate conceptions). These mechanisms most frequently explained the "hidden" sequences of evaporation. While students articulated in both their pre- and postunit models that temperature change was necessary for water to 'leave clouds' and gravity was necessary for water to 'fall to the ground', they also relied on and incorporated alternate conceptions for *how* and *why* water returns to the sky from the clouds. Most frequently students' sequences of water vapor returning to the sky were drawn as one large quantity moving as a single entity from

the ground surface to a large dark cloud. Their mechanism for this movement was represented and articulated as tubes reaching from the ground to a specific cloud. For example, while Caroline represented that an increase in temperature was the mechanism for liquid water to change its state to water vapor, her mechanism for water vapor returning to the sky was "invisible helium tubes" that "take the water vapor and bring it up to the cloud" (Figure 1, P54:037). Additionally we found some students were unable to conceptualize how water might return to the sky so they did not include sequences for water's return. They articulated mechanisms for water to remain on the because the Earth 'holds' water by "sucking" or "absorbing" water to store (N.EM2, W.S1M1).

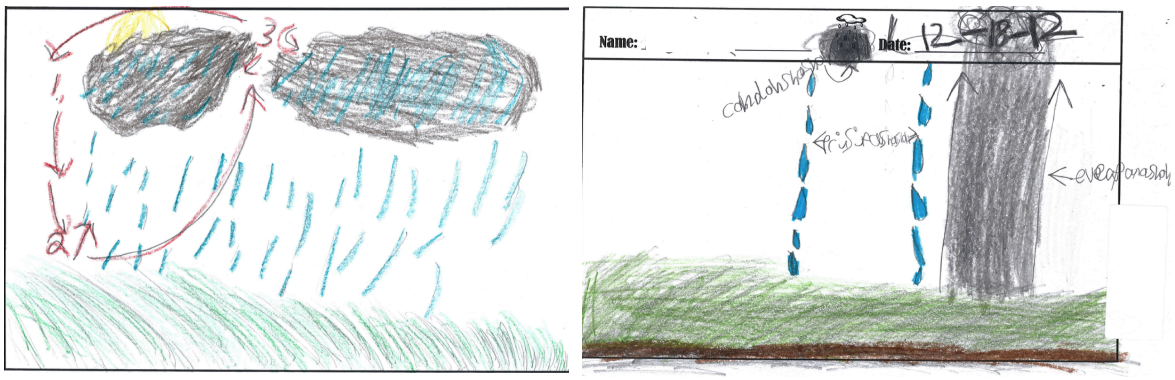


Figure 1. Student model of 'hidden sun' (S9:N) and 'evaporation' (W:S17)

Conclusions and Implications

Scientific modeling is a core scientific practice highlighted in the *Next Generation Science Standards* (NRC, 2012). Modeling practices, however, remain underemphasized in K-12 science, particularly in the elementary grades (Windschitl, Thompson, & Braaten, 2008; NRC, 2007) and, as a result, little research exists to guide efforts to foster epistemically-rich, model-centric elementary science learning environments. This study leverages and makes contributions to research on model-based science teaching and learning (Lehrer & Schauble, 2004; Schwarz et al., 2009; Windschitl, Thompson, & Braaten, 2008) and content- and practice-based learning progressions (Alonzo & Steedle, 2008; Lee & Liu, 2010; Mohan, Chen, & Anderson, 2009; Stevens, Delgado, & Krajcik, 2010) through the development of a learning performances framework and empirical findings from its use to investigate 3rd-grade students' model-based explanations for hydrologic cycling.

Past research has illustrated aspects of the water cycle, a core subject in the K-12 curriculum (NRC, 2012), that are often challenging for students (Bar, 1989; Henriques, 2002). Though recent research has shown that elementary students can learn to effectively engage in scientific practices, including scientific modeling, to develop conceptual understanding (Hapgood, Magnusson, & Palinscar, 2004; Hardy, Jonen, Möller, & Stern, 2006; Herrenkohl & Cornelius, 2013; Lehrer & Schauble, 2006; Manz, 2012; Metz, 2004; McNeill, 2011), little work has been conducted to investigate students' formulation of model-based expalanations for hydrologic cycling. Findings from this study highlight the range of ideas evident in students' model-based explanations for system processes that underlie the water cycle. Results suggest 3rd-grade students emphasize sequences of water movement and statements that map representations of water movement onto real-world phenomena more effectively than for the forms of water. Further, students do not foreground the relationship between water and the Earth in their model-based reasoning about the water cycle. These findings provide insight not only into leverage points through which to foster early learners' formulation of model-based explanations for water, but also those concepts and epistemic features for which curricular and instructional guidance would likely be most impactful. Further work is needed to explore how to build upon the domain-specific (i.e., hydrologic cycle) epistemic commitments and conceptual strengths students exhibit, as well as appropriate design and implementation of scaffolds for students' use of models to reason about the water cycle.

This work also highlights the utility of a set of *empirically-based learning performances* (Krajcik, McNeill, & Reiser, 2007), developed from the ground up as part of a design-based research program, that can be used to both design and study discipline-specific model-centric science learning environments. The integrated learning performances framework developed here illustrates critical trends in 3rd-grade students' model-based explanations for hydrologic cycling, such as a de-emphasis on unobservable components of water systems, bidirectionality of system processes, and rationales for representational norms. Measureable levels of model-based explanations for water cycle processes, based upon empirically-derived trends in students' thinking, have grounded the ongoing study of the curricular and instructional components of these 3rd-grade science learning environments. Such work is necessary to inform the future development of learning progressions that a) integrate the scientific practices of modeling AND domain-specific concepts and b) are sufficiently robust to account for elementary students' learning within the Earth Sciences, a domain within which little learning progressions work has thus far been carried out. This learning performances framework, as well as the

modeling task and findings from project research, will inform ongoing efforts to design elementary science learning environments through the development of science curriculum materials and assessment resources, as well as efforts to foster teachers' instructional practices that promote ALL students' model-based learning.

References

Alonzo, A. C. & Steedle, J. T. (2008). Developing and assessing a force and motion learning progression. *Science Education*, 93(3), 389-421.

American Association for the Advancement of Science. (2007). *Atlas for scientific literacy*. Washington, DC: National Science Teachers Association Press.

Bar, V. (1989). Childrens' views about the water cycle. *Science Education, 73*(4), 481-500.

Braaten, M. & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education, 95*(4), 639-669.

Hapgood, S., Magnusson, S., & Palinscar, A. (2004). Teacher, text, and experience: A case of young children's scientific inquiry. *The Journal of the Learning Sciences, 13*(4), 455-505.

Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*(2), 307–325.

Henriques, L. (2002). Childrens' ideas about weather: A review of the literature. *School Science and Mathematics, 102*(5), 202-215.

Herrenkohl, L. & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. *The Journal of the Learning Sciences, 22*(3), 413-461.

Krajcik, J., McNeill, K. L., & Reiser, B. (2007). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*(1), 1-32.

Lee, H.-S. & Liu, O. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education, 94*(4), 665-688.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 371-388). Cambridge, MA: Cambridge U. Press.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models*: Second edition. Cary, NC: SAS Institute, Inc.

Manz, E. (2012) Understanding the codevelopment of modeling practice and ecological knowledge. *Science Education, 96*(6), 1071-1105.

McNeill, K. L. (2011). Elementary students views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching, 48*(7), 793-823.

Metz, K. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design. *Cognition and Instruction, 22*(2), 219-290.

Mohan, L., Chen, J., & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46*(6), 675-698.

National Research Council. (2012). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: National Academies Press.

National Research Council. (2007). *Taking science to school: Leaning and teaching science in grades K-8*. Washington, DC: The National Academies Press.

Patton, M. Q. (2001). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Schwarz, C., Reiser, B., Davis, E. A., Kenyon, L., Acher, A., Fortus, D., Schwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-654.

Stevens, S.Y., Delgado, C., & Krajcik, J.S. (2010).Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching, 47*(6), 687-715.

Shin, N., Stevens, S.Y., & Krajcik, J. (2010). Designing assessments to track student progress. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences* (ICLS 2010). University of Illinois at Chicago: Chicago, IL.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education, 92*(5), 941-967.

Acknowledgments

This material is based upon work supported by the National Science Foundation (DRL-1427115). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF.

Problematizing as Scaffold for Engaging in Scientific Argumentation

Mon-Lin Monica Ko, Learning Sciences Research Institute, University of Illinois at Chicago, mlko@uic.edu
1240 W. Harrison Street Room 1570F Chicago, IL. 60607

Abstract: Supporting students' engagement in disciplinary practice is a major tenant of recent reforms in science policy and reform, and yet eliciting and building upon students' everyday knowledge through engagement in scientific practice remains a central challenge to this work. This paper presents evidence of how *problematizing* – a term previously used to design scaffolds embedded in technological tools and supports utilized in smaller, face-to-face settings – can work in a class-wide setting, as a support for noticing gaps in students' current explanations both while motivating the need to engage in investigations and while building consensus explanations to account for science phenomena.

Arguing to Learn: Generating Explanatory Accounts for Phenomena

A focus of reforms in standards, learning environments, teacher preparation programs and professional development is to support teachers' and students' engagement with scientific practices to generate explanations for real-world phenomena (Achieve, 2013; Windschitl, Thompson, & Braaten, 2008). Engaging in argument as a means for developing scientific explanations has been characterized as *arguing to learn* because arguments are generated to explain phenomena and build an understanding of disciplinary core ideas (McNeill, Lizotte, Krajcik, & Marx, 2006). Classrooms in the U.S. rarely engage in arguing as a means of constructing evidence-based explanations (Pasley, Weiss, Shimkus, & Smith, 2004), and thus, scaffolds may be needed to enculturate students into the practice of articulating claims, evidence, and reasoning in their explanatory accounts (Berland & McNeill, 2010). Developing evidence-based accounts for phenomena can be challenging when multiple candidate explanations can be constructed using the same evidence, when similarities or differences between alternative explanations are not obvious, or in cases where *how* the explanation accounts for the to-be-explained phenomena is unclear. Because students frequently hold conflicting ideas about science phenomena, identifying the differences between multiple candidate explanations may support students in building on or integrating explanations generated from their everyday knowledge of the world and evidence-based accounts (Linn & Eylon, 2011). These challenges point to the need for instructional contexts and scaffolds that support students in becoming increasingly adept at generating evidence-based arguments to explain phenomena.

Problematizing as Discursive Scaffold for Complex Learning

Problematizing has been discussed in the Learning Sciences literature as a key principle for fostering productive disciplinary engagement and also as a design principle for software tools to support complex learning (Engle & Conant, 2002; Reiser, 2004). In contrast to simplifying or making complex tasks more explicit, problematizing maintains complexity while recruiting students' attention, resources to generate dissonance and curiosity for resolving the task at hand to support students in engaging with difficult problems (Reiser, 2004). This paper makes the case that problematizing scaffolds can support students' engagement in disciplinary discourse practices and build criteria for generating evidence-based explanations for phenomena.

Why might problematizing be a scaffold for engaging students in arguing to learn? For one, problematizing may help students articulate their reasoning and thinking processes, a key aspect of engaging in scientific argumentation (Osborne & Patterson, 2011). Second, when multiple explanations are plausible, argumentation can be used to identify differences between alternative explanations and make clear how evidence and reasoning supports or refutes these candidates. Problematizing may thus support this decision-making process by helping students categorize or distinguish between alternative accounts. Lastly, problematizing students' existing ideas can bring gaps and disagreements to light. Uncovering differences between alternative explanations or gaps in one's reasoning is crucial to the construction of well-grounded arguments supported by sound evidence and reasoning and also takes into account and rebuts alternatives. Though features of this type of scaffold appear to be well suited to supporting the established challenges of engaging in argumentative discourse, how this can be accomplished has not been well articulated.

Moreover, promoting conceptual change involves the disciplinary practice of scrutinizing and revising one's conceptual understandings in light of new evidence, (Duschl, 2000). Problematizing scaffolds can help make explicit the differences between candidate explanations generated from everyday knowledge or from first-hand evidence and established scientific principles. Engaging students in sense-making discussions in which they are accountable to their peers, standards of reasoning, and disciplinary norms for knowledge building can support students' constructing, evaluating and critiquing scientific explanations (Michaels, O'Connor, & Resnick, 2008). Understanding how teachers and students establish and instantiate a disciplinary standard of

reasoning and arguing to develop explanations for phenomena is critical. Sense-making talk holds great potential for problematizing students' everyday accounts as insufficient explanations, and generates a need to utilize evidence and science principles to generate a coherent explanatory account.

In this paper, I draw on 2 cases from one 6th grade classroom to illustrate discursive practices that mirror the core principles of problematizing as a scaffold for productive engagement in science practice. I first describe the larger longitudinal study that documented variations in how four 6th grade classrooms took up the science practices. Then, I specify how one teacher's classroom was identified as a particular case of *problematizing,* and finally, elaborate on two cases from this class to illustrate how problematizing scaffolded the process of distinguishing and integrating everyday and evidence-based explanations to motivate the need to engage in investigations, and also to support students to coming to a consensus explanation.

Methods

Curriculum Context

This study took place in the development and national field trial testing of reform-based science curriculum, *Investigating and Questioning our World Through Science and Technology* (IQWST). The 6th – 8th grade curriculum engages students in scientific practices to investigate core ideas in science (Krajcik, McNeill, & Reiser, 2008). IQWST lessons are comprised of investigations that involve students in asking and observing science phenomena, conducting data investigations, reading relevant texts, and discussions to support students' sense making of core disciplinary ideas and practices. Drawing on the work of pilot studies, which uncovered that teachers' curricular adaptations varied in their intent to support students' sense making of the disciplinary core ideas (Ko & Reiser, 2010), 3 multi-day lessons, spanning the Physics and Chemistry units, were selected to better understand how teachers and students engaged in constructing investigation questions, data collection, and the development of a consensus explanation for various phenomena. The 3 multi-day lessons were observed throughout the school year in four 6th grade classrooms, spanning from October to February during 2010-2011 to provide insight into teachers' disciplinary practices for students who have not yet had any experience with the IQWST curriculum. The two cases presented in this study come from the first 6th grade IQWST unit on light. In both Physics lesson 6 and lesson 11, students are engaged in the work of developing and refining their models of how light interacts with matter: exploring why light behaves differently when interacting with a mirror or a piece of paper, and what happens to white light when it interacts with colored filters.

Analytic Processes

The larger longitudinal study investigated variations in how classrooms engaged in scientific practices. To first describe how classrooms engaged in asking questions, conducting investigations and building explanations, all lessons observed across the four classrooms were videotaped and transcribed. Ongoing field notes of the classroom activity and pre and post interviews with the participating teachers supplemented the videotaped classroom observations as resources for developing and refining candidate conjectures. Field notes were taken for each day a lesson was enacted. A finalized version of the field notes, documenting significant events and their time stamps in classroom observations were recorded and finalized within 1-2 days of the observation. Ongoing memos and jottings, comprised of reviews, reflections, and ongoing questions that emerged from the classroom observations and interviews, were used to further refine ongoing hypotheses about the variations in how classrooms engaged in scientific practices, as well as the factors that contributed to the differences across the four classrooms.

As a result of this ongoing analysis, I found that although classrooms engaged in the same set of data collection investigations, *how* these investigations were motivated and subsequently *made sense of* differed across the four classrooms. This observation led to an increasing focus on the sense-making discussions that preceded and followed data collection activities across the four classrooms.

Discourse Analysis

To capture variations in how classrooms motivated and made sense of investigations to generate explanations, classroom talk was divided into episodes, based on the scientific practices in which students were engaged. Using a grounded approach of moving back and forth between the findings from prior work, existing literature on classroom discourse and scientific practices, and looking for confirming and disconfirming evidence within the available data, 22 codes were developed to describe ways in which students and teachers were engaged in scientific practices. These codes captured how classrooms went about asking questions, making sense of data, and generating explanations as well as the cognitive and material resources that were used to do this work. A second coder coded 49% of the total data corpus, with a reliability of $\kappa = .82$. Agreement was reached through discussion on all discrepant episodes.

Looking for patterns in the resources that teachers and students drew on to build explanations for phenomena revealed the epistemic criteria utilized to do knowledge-building work. Within the entire set of

codes identified and refined through the analysis of the total data corpus, 3 codes captured the types of resources with which teachers and students construct, critique or defend candidate explanations for scientific phenomena: 1) *data from classroom investigations*, 2) *previously established principles and models*, and 3) *students' everyday knowledge*. 'Data' consisted of the qualitative and quantitative observations that students generated during classroom investigations for the current lesson. 'Previously established principles and models' were defined by the evidence-based claims generated through previous investigation activities and sense-making talk that addressed each unit's driving question; these principles were displayed in the class as visual representations or written statements and were co-constructed by students and teachers. Finally, 'everyday knowledge' – derived from student's daily observations of the natural world, was also used to generate candidate explanations.

Teacher and Case Selection

The coding process revealed variations in how often and in what ways the aforementioned resources were used to construct candidate explanations differed across classrooms and over the course of the 3 observed lessons. While all 4 classrooms utilized the 3 resources during discussions, two classes used *previously established principles* and *data from investigations* most often to generate questions and to build explanatory accounts for how light interacts with matter. The remaining classrooms put forth a greater proportion of explanations drawn from students' everyday knowledge, although the classrooms conducted identical investigations. These patterns suggested that investigations were treated differently across the four classrooms; in some, explicitly as opportunities for students to build on their existing knowledge with science principles and data to generate explanatory accounts for phenomena.

I reviewed the transcripts of all four teachers, and then zeroed in on Laura's class, to better understand how this was accomplished. In her class, candidate explanations generated from data, students' everyday knowledge, and previously established principles were juxtaposed to *motivate a need* to engage in investigations, but also as a means to *constructing consensus explanations* when alternative accounts were present. This is in contrast to classrooms where investigations goals for investigation were stated by the teacher, rather than driven by surfacing gaps in students' current explanations. During consensus-building discussions, Laura's class went through a laborious process of presenting alternative explanations and identifying evidence for or against candidate explanations to arrive at a single consensus explanation. Other classrooms, in contrast, simply identified and re-stated the same explanation, and often without accounting for or refuting alternatives.

After identifying these patterns, revisiting transcripts, and returning to extant literature on argumentation and explanation, I conjectured that these patterns revealed a unique case of how *problematizing* scaffolded students' engagement in scientific practice. A review of existing literature on argumentation in science classrooms further confirmed this conjecture; using frameworks specifying progressions in students' argument patterns and products (e.g. Berland & McNeill, 2010; Osborne & Patterson, 2011), the talk exhibited in Laura's classroom demonstrated sophisticated argumentation when motivating investigations and generating explanatory accounts for phenomena. I identified 2 such cases within the transcripts from Laura's class, presented in detail below, to highlight *how* problematizing is accomplished, *what* is problematized, and the key *scientific practices* (Achieve, 2013) that these discourse patterns enable students to engage with. Thus, these episodes from Laura's class were selected as *paradigmatic* cases (Flyvbjerg, 2006) of problematizing as scaffold. Analysis of students' performance on several pre post measures, when compared to other classrooms, also indicated greater gains in these students' ability to attend to disciplinary criteria when generating science explanations (Buckingham & Ko, 2013). For this reason, I selected segments of Laura's classroom as a rich case of how students are scaffolded into arguing to learn. Laura has over 20 years of experience teaching both science and math, and holds a degree in elementary education. Approximately 10 hours of video footage was observed in Laura's classroom from October to February throughout the study. At the time of the study, Laura used preliminary version of the field trial IQWST curriculum for 5 years, conducted professional development for new teachers, and worked closely with one of the lead universities creating the curriculum materials. The 3 other teachers were in their 2nd year of using the curriculum materials.

Findings

Case 1: Problematizing to Motivate Investigations

In Physics lesson 6, students investigate how light interacts with various types of objects. In previous lessons, students have already begun to describe how light interacts with matter by describing that light "bounces" off. However, the class has yet to explore how this bouncing differs when light is shone on different types of objects, such as a piece of paper or a mirror. Though this phenomenon may seem quite commonplace, everyday accounts for why light bounces differently are examined in Laura's class as worthy of further inquiry.

The lesson begins as Laura takes a flashlight and shines light onto a piece of paper that is taped to the whiteboard in the front of the class. While the students readily draw on previously established principles to explain that light bounces off the paper, students have difficulty explaining why the light bounces differently off

a mirror. Laura relentlessly probes students to explain why light interacts with the two objects in such different ways (claims of how the light interacts with objects underlined below, the teacher's probing questions *italicized*, and description of non-verbal moves or behavior in (parentheses)):

Episode	Resources used	Teacher and student talk
1	Previously established principles and models	Laura: So the light's hitting the paper. Are you able to see the paper? Students: Yes Laura: Okay why are you able to see the paper? (Almost all students have hands raised) Oh look at all these hands. Why are you able to see the paper Joshua? Joshua: Because the light from the light source is bouncing off the paper and into my eye. …(Additional talk ensues)…
2	Data from investigation	Laura: So light's just hitting the paper and bouncing off and going into your eye, and now what if I use a mirror (Teacher shines light on a hand held mirror, directing light back and forth towards the students, as students begin to yell "Ah my eyes!")…*When I did the paper you weren't yellin*g (Teacher aims the light back at the paper) Matt: Because it's bouncing off differently Laura: *So what's makes it different when light bounces* (Teacher continues to use the mirror to shine light on the students)…So what happens? *Why do you think it's different* from when the light hits your paper? And when it hits the mirror? Alan?

Although the Laura's students use previously established principle that light comes from a source, bounces off the paper and into the eye, to explain how light interacts with paper (episode 1), this principle comes under scrutiny after Matt admits that the light is "bouncing differently" off of the mirror. Laura underscores this by directing the light in various directions (at times hitting the students' eyes) to further emphasize the contrasting ways that light interacts with the mirror and paper (episode 2). Building off of Matt's claim that light bounces differently, Laura asks the class, "so what makes it different…why do you think it's different?" The previously established principle "light bounces off…into my eyes" no longer provides a coherent account of how light interacts with both paper and mirrors, signaling a need to revise their previously established principles and models.

In response to Laura's probing, a series of exchanges follows in which students attempt to generate an explanation for these differences in how light bounces. As students begin offering candidate explanations drawing on their everyday knowledge, Laura utilizes previously established principles and observational data to further reinforce the insufficiency of these candidate explanations. The juxtaposition of previously established principles and models of how light interacts with matter and the observed difference in how light interacts with paper and mirrors generates a momentum and need for students to engage in further investigation:

Episode	Resource used	Teacher and student talk
3	Students' everyday knowledge	Alan: Uh I think its different 'cuz when it hits the paper it's carrying a - there's kind of photographic message like when it's hitting the mirror (Inaudible) light rays so there's no picture Laura: *Okay so what would you say? (*Looking at another student, not Alan)...do you have any ideas? Patrick: I think - (Inaudible) maybe its 'cuz (Shrugs his shoulders)
4	Previously established principles and models	Laura: I mean you said the light was bouncing off paper and going into your eyes (uses the mirror again playfully to shine into Ss eyes) we'll just assume that that happens. Okay – Irene?
5	Students' everyday knowledge	Irene: when you use paper, it's just shining (Inaudible) but when it hits the mirror its bouncing off the mirror (Inaudible) Laura: *So why does it do it for a mirror and not for a paper*? (Lots of student hands raised in response to this question) Irene: 'Cuz the mirror is like if we look you can see your reflection...(Inaudible) Teacher: *Why* (asking Irene, who appears to be stumped)? Good questions.
6	Data from investigations	Laura: So *is* the light hitting different objects differently? Students: Yes. Laura: *Why*?...How does light know how to do that?...*should we collect some more information about this*?

In the above episodes, students generate several different candidate explanations: Alan proposes that the paper projects a photographic message of light (episode 3) and Irene suggests that light *shines* on paper but *bounces* off a mirror (episode 5). In both instances, Laura probes these candidates by asking "why?" or "what would you say?" and reminds students, "You said the light was bouncing off…and going into your eyes" (episodes 4 and 5). Contrasting previously established principles and models from students' prior investigations in the unit and the observations of light shining on paper and mirrors further underscore the paucity of their existing explanations to account for the differences between how light reflects off the mirror and paper (episode 6).

As a result of these series of exchanges in which the teacher and students draw on data, students' everyday knowledge, and previously established principles, discrepancies between students' previously established principles and models and their current observations of how light interacts with mirrors and paper are highlighted. By making the gap between what students see and what they know explicit, the upcoming investigation becomes an opportunity to collect data to support a more complete explanatory account. These set of exchanges in Laura's classroom establishes a disciplinary-specific purpose to what might otherwise be viewed as a science lab activity, and creates a demand for students to revise their existing model in light of these perplexing observations of how light interacts with objects.

Case 2: Problematizing to Build Consensus Explanations

In Physics lesson 11, students work with C-spectra, colored filters, and an overhead (emitting white light) to explore components of white light and how white light interacts with filters. At this point in the Physics unit, students have begun developing and refining a model of how light interacts with matter, coming to consensus through investigations and discussions that white light interacts with objects by *reflecting* off, *transmitting* through, or getting *absorbed* by the object. During this lesson, students try to extend this existing model to explain colored light and how white light interacts with colored filters.

Laura begins this inquiry by projecting light through an overhead projector covered by a piece of paper with a rectangular slit cut in the middle, covered by C-spectra paper (which breaks white light into its component color spectrum). Without any colored filters, the white light transmits through the spectra paper and projects a rainbow of colors representing the components that makes up white light. However, when the orange filter is placed over the rectangular slit and C-spectra, the blue and purple components of white light seem to "disappear". After the class works together to investigate the colors of light "disappear" by using various colors filters, Laura challenges students to generate an explanatory account for their observations, asking, (specifically about the effect of the orange filter on white light), "Where are those colors…where is the blue and purple light going?" (Line 519-523, 11/15/2010)

In response Laura's request, students begin to put forth candidate explanations for the missing light. In the transcript below, Jonathan and Lana offer up candidate explanations, drawn from their everyday knowledge and previously established principles, to account for the missing blue and purple light. An issue arises when Lana misappropriates a previously established principle to explain the missing light; while Laura acknowledges the validity of the principle, she points out this principle does not adequately explain how the white light from the overhead (composed of visible spectrum, *including* blue and purple light) hits the colored filter and somehow disappears as it gets transmitted to the whiteboard:

Episode	Resource used	Teacher and student talk
1	Students' everyday knowledge	Jonathan: <u>It's disappeared.</u> Laura: It's disappeared, but I'm going to tell you I did not do any magic…*this is going to be a phenomenon that we can explain*…so, Lana, what do you think happened to the color?
2	Previously established principles and models	Lana: <u>You don't need the blue to make orange</u>, <u>so I don't think it shows up</u>. And because you need blue to make purple, I don't think it shows up either.
3	Data from investigation	Laura: Okay. So you're right. So we don't need it. *However, you didn't answer my question.* So here's the white. It disappeared. Where did it go? Here's the white light (pointing to the light bulb in the overhead). It's coming up. It gets to here (Teacher points to the projection lens above the overhead screen). Do you all agree it gets to here? Several students: Yeah. Laura: *So where does it go?*

After Jonathan suggested that the light has disappeared, Laura pushed students to account for the underlying mechanisms of how this has happened by saying, "this is a phenomenon we can *explain*" (episode 1). Using a prior activity in which she mixed various colors of light, Lana used principles of color composition, specifically,

that *orange light is not made up of blue and purple light*, to explain that the light does not show up (episode 2). In response to Lana's candidate explanation, Laura returns to the demonstration with the overhead projector, pointing out that while Lana's right - blue light is not needed to create orange light, the white light (containing the blue and purple spectrum) travelled through the projector, and interacts with the C-spectra paper and the orange filter. By establishing consensus that blue and purple light were projected from the overhead, Laura simultaneously points out the shortcomings of Lana's explanation while requesting further explanation from her classmates, asking "so where does it go?" (Episode 3)

Following this critique of Lana's candidate explanation, Dale and Caitlyn draw on their everyday knowledge to generate additional candidate explanations for the missing light. When Laura later returns to Lana a second time for an explanation of the observed phenomena, Lana now utilizes the language from previously established principles about how light interacts with matter to suggest that the filter might be *reflecting* the missing light. Laura offers Lana's idea to the class, and then supports them in drawing on previously established principles and observational data to determine if there is sufficient empirical evidence in support of Lana's claim. Through the following exchanges, the class moves from generating candidate explanations to using principles to find disconfirming evidence to support a subset of these candidates:

Episode	Resource used	Teacher and student talk
4	Students' everyday knowledge	Dale: It gets re-filtered back. Laura: Okay…Caitlyn? Caitlyn: I think it's still there, but the orange filter's…covering it.
5	Previously established principles and models	Laura: [The] Filter's covering it. So it's covering it so we can't see it, right? And, sorry, Lana, tell me your answer again. Lana: The filter reflects it because you don't need blue and purple to make orange.
6	Data from investigation	Laura: Okay. *So what do we think about [the filter] reflecting it*? Matt: I don't know. Laura: *It's reflecting which colors, Lana*? Lana: Blue. Laura: *Does anybody see blue?* Students: No. Kyle: No. I see orange in the background.
7	Previously established principles and models	Laura: *If it was going to reflect blue, would you see blue*? Kyle: Yes. Mike: I don't know. Laura: So we could still go with what Colleen's saying, right? Like it gets there, but then we don't know where--it's going somewhere. Assad: It's absorbed. Laura: Assad? Assad: It's absorbed. Laura: Why do you think absorbed? Assad: Well, because it's not being reflected and it's not being transmitted it's not being scattered.

After Dale and Caitlyn offer up their ideas, Laura returns to Lana for further elaboration on a previously stated explanation. Although her stated idea (the filter reflects the missing light) differs this time, her reasoning for "because you don't need blue and purple to make orange" remains the same. Laura publicly acknowledge this claim and asked students evaluate it by asking "so what do we think about it…?" (Episode 4-5) Lana's explanation that the missing light is *reflected* is taken up and evaluated in episode 6, where Laura taps students previously established principle that when light reflects, *it bounces off objects and into our eyes*. After students realize that they do not see blue light being reflected off the orange filter, Laura further emphasizes the implications of this observation based on their principles, saying, "If it was going to reflect blue, would you see blue?" The combined effort of first gathering additional (observational) data and the use of previously established principles leads Assad to deduce that the missing light is absorbed by the filter because "it's not reflected…transmitted…scattered" (in episode 7).

Over the course of the 7 episodes presented above, Laura's class not only utilizes variety of resources to generate candidate explanations, they also critically evaluate them using previously established principles to identify the relevant disconfirming evidence. Although the teacher instigates the critique of candidate explanations, students are involved in the joint effort to identify contradictory evidence and offer alternative explanations that adhere to their previously established principles. After Assad proposes his candidate explanation, the classroom continues to work with several colored filters to gather additional supporting

evidence for this candidate explanation, bolstering its validity. Later, the class deepens the idea that white light interacts with colored filters by *absorbing* a subcomponent of its total spectrum while letting others *transmit* through by using this same explanation to explain how colored sunglasses work. Through this sense-making talk, the students eliminated alternative explanations using evidentiary support and increased the generalizability of their current explanation – both critical aspects of epistemic criteria for knowledge building in science.

Discussion

The two cases presented in this paper, both from Laura's class, exemplify how sense making discussions serve to problematize students' understanding, thereby engaging students in sophisticated argumentative practices. The first case demonstrates that when discrepancies between students' current understanding and the observed phenomena are highlighted, students are motivated to engage in investigations to generate a more satisfactory explanatory account. The second case illustrates that by drawing attention to evidentiary support for candidate explanations, students attend to the fit between evidence, alternative explanations, and the disciplinary processes for moving towards consensus. In both of these cases, the problem-space is not simplified; rather, students are encouraged to offer up a multitude of candidate explanations and supported in sorting through these ideas (see Table 1 for summary of 2 cases).

Table 1: Problematizing as scaffold for motivating investigations and generating explanations

	Case 1: Problematizing to Motivate Investigations (Physics 6)	Case 2: Problematizing to Build Consensus Explanations (Physics 11)
What resources are used to problematize?	Data from investigations (observations) Students' everyday knowledge Previously established principles and models	Data from investigations (observations) Students' everyday knowledge Previously established principles and models
Are discrepancies highlighted?	Yes – between students' observations and previously established principles	Yes – between candidate explanations and available evidence
Are alternative explanations presented and tested?	Yes – class puts forth several alternatives but they are found to be insufficient based on students' observations	Yes – class draws on data and previously established principles to eliminate candidate explanations
Scientific practice students engage with	Evaluate limitation of existing model or explanation; analyze and interpret data to provide evidence for phenomena	Construct explanations with evidence; using scientific reasoning to show how evidence supports or refutes explanations
Function of teacher's discourse moves	Documenting candidate explanations publicly and marking explanations as objects of attention; requiring students to reason and justify candidate explanations; drawing attention to the fit between claims and available evidence.	

Based on existing literature, the patterns of talking in the cases presented here embody the core principles behind how problematizing scaffolds students into the practice of building causal accounts for perplexing phenomena. Whether the class is *motivating* investigations or *building consensus explanations*, through the juxtaposition of data, students' everyday knowledge, and previously established principles and models, students are pushed to coordinate this body of knowledge by critiquing and evaluating gaps in their current understanding and to use principles and data as evidentiary support. As a result, students begin to see investigations as opportunities to build stronger explanatory accounts and view the work of constructing consensus explanations as the process of coordinating claims, evidence, and reasoning for multiple accounts.

It is important to note that these two cases were not selected as examples of problematizing scaffolds at work from the outset of the study. Taking this analytical lens emerged through a thorough analysis of the discourse from all four 6th grade classes, and by returning to the existing literature to describe the type of scaffold that was utilized in Laura's class. Notably, students drew on in similar everyday knowledge, previously established principles and models, and data from investigations in all of the 4 observed classrooms. It was not simply drawing on these bodies of knowledge, but rather, identifying gaps and insufficiencies between them that moved the students forward in Laura's class. Identifying these episodes as cases of problematizing emphasizes the role of scaffolds that go beyond a prescriptive understanding of science argument as claims, evidences, and warrants. We see students building criteria for evaluating claims organically out of reflecting on the insufficiency of one's existing knowledge and understanding the role of science principles and data as resources for constructing increasingly sufficient explanations in Laura's classroom. The remaining 3 classrooms did not exhibit the consistent tensions born out of noting the explanatory power of different candidate explanations, or how these explanations differentially accounted for the observed phenomena as was observed in Laura's class.

Several follow up questions are worthy of further inquiry. Are these scaffolds taken up within students' discursive repertoire over time? That is, does Laura take on less of the responsibility, while students take increasing ownership over critiquing and evaluating one another's explanatory claims? Moreover, does this form of critical and evaluative talk become an integral and spontaneous component of the classroom practice? Identifying additional exemplars of problematizing scaffolds can contribute to the design of social, discursive, and technological tools in ways that do not remove the complexity of the knowledge building work in which students engage (Davis & Miyake, 2004). Furthermore, there is a need to identify other such pedagogical moves that constitute a problematizing repertoire for teachers.

Lastly, the examples presented from Laura's class provide insight into how to apprentice students into the scientific practices of arguing to learn, and specifically, integrating evidence-based and everyday explanations for phenomena. Rather than discounting the everyday accounts that students bring into the classroom, or relying solely on the data generated from investigations as resources for knowledge building, understanding the disjuncture and connections between these bodies of knowledge can equip teachers with the necessary probes that encourages students to add, distinguish, and sort out competing candidate explanations using evidence and science principles (Linn & Eylon, 2011).

References

Achieve. (2013). Next Generation Science Standards. Retrieved from http://www.nextgenscience.org/next-generation-science-standards

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765–793.

Buckingham, B.L.E, Ko, M. (2013). *Using Classroom Discourse to Account for Differences in the Causal Coherency of Written Explanations*. Paper presented at the National Association for Research in Science Teaching, Rio Grande, Puerto Rico, April 6 – 9.

Davis, E. A., & Miyake, N. (2004). Explorations of scaffolding in complex classroom systems. *The Journal of the Learning Sciences*, *13*(3), 265–272.

Duschl, R. (2000). Making the nature of science explicit. In *Improving science education: The contribution of research* (pp. 187–206). Open University Press.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, *20*(4), 399–483.

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 219–245.

Ko, M. & Reiser, B. J. (2010). *Connections, clarity and logistics: science teacher's enactments of reform-based science curricula.* Paper presented at the American Education Research Association, Denver, CO, April 30 – May 4.

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-Goals-Driven Design Model: Developing Curriculum Materials that Align with National Standards and Incorporate Project-Based Pedagogy. *Science Education*, *92*(1), 32.

Linn, M. C., & Eylon, B.-S. (2011). *Science Learning and Instruction: Taking Advantage of Technology to Promote Knowledge Integration* (1st ed.). Routledge.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *Journal of the Learning Sciences*, *15*(2), 153–191.

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297.

Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, *95*(4), 627–638.

Pasley, J. D., Weiss, I. R., Shimkus, E. S., & Smith, P. S. (2004). Looking inside the classroom: Science teaching in the United States. *Science Educator*, *13*(1), 1–12.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 273–304.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*(5), 941–967.

Acknowledgments

The research reported here was supported by the National Science Foundation through Grant ESI-0227557, ESI-0439493, ESI-0101780 and DRL 1020316. The opinions expressed are those of the author and do not represent views of the National Science Foundation.

Three Diagnoses of Why Transfer Across Disciplines Can Fail and Their Implications for Interdisciplinary Education

Eric Kuo, Stanford University, Stanford, CA 94305, erickuo@stanford.edu
Danielle Champney, California Polytechnic State University, San Luis Obispo, CA 93407, dchampne@calpoly.edu

Abstract: One goal of interdisciplinary educational efforts is to increase students' ability to transfer knowledge from one disciplinary context to another. The approach taken to foster this transfer depends on the diagnosis of why this transfer can fail in the first place. Although a common diagnosis focuses on content knowledge and problem features as explanatory, there exist other, less prevalent diagnoses for why transfer across disciplines fails. In this paper, we show how one student responds differently to two similar problems set in physics and calculus problem contexts. We argue that beyond a *content knowledge diagnosis*, an *epistemology diagnosis* and an *accountability diagnosis* can also plausibly contribute to an explanation of why this student approaches these two similar problems differently, presenting additional considerations in fostering interdisciplinary transfer.

How Approaches to Interdisciplinary Education Imply Diagnoses of Why Transfer Fails

The transfer of content knowledge across different disciplines is one goal of a multidisciplinary education system. A typical curriculum for an undergraduate STEM major is designed with the expectation that students can and will, in any particular course, apply knowledge and skills learned in other disciplinary courses – for example, the typical undergraduate physics curriculum is designed expecting that students can and will transfer knowledge learned in prerequisite math courses into those physics courses; similarly, engineering courses expect students can transfer-in knowledge from math and physics courses.

Although these various courses are largely designed independently from one another, some interdisciplinary course reform efforts have developed, aiming to bolster transfer across the disciplines. The most common interdisciplinary approach is to align the content between two courses in order to construct content connections across the disciplines (e.g. Al-Holou et al., 1999; Dunn & Barbanel, 2000; Plomer, Jessen, Rangelov, & Meyer, 2010). Common goals of these efforts include avoiding haphazard coverage of related content across disciplinary courses as well as highlighting content connections between disciplines to decrease compartmentalization of knowledge and encourage future transfer to different disciplinary problem contexts.

Approaches such as these reveal (either explicit or implicit) views that poor content knowledge generalization and problem-specific surface features are explanatory in understanding why transfer across disciplines can fail. This *content knowledge diagnosis* is aligned with classical views of transfer that found reduced spontaneous transfer of a learned problem solution when the surface features of transfer problems were made more dissimilar from the initial training cases (Gick & Holyoak, 1980; Reed, Ernst, & Banerji, 1974). Such transfer is improved when multiple training cases were used, supporting the development of a more general solution schema not tied to the surface features of any one particular case (Gick & Holyoak, 1983; Reeves & Weisberg, 1994). In the same way, the alignment and co-presentation of content in interdisciplinary efforts aims, in part, to prevent the content knowledge from being tied to any one disciplinary context while also providing practice in applying that content knowledge in multiple disciplinary problem contexts.

We argue that there exist other possible diagnoses for why transfer across disciplines fails, which are not as prevalent as the *content knowledge diagnosis*. In this paper, we present one student, Will, who reasons differently on two similar problems set in physics and calculus disciplinary contexts. Investigating why his reasoning differs on these problems leads to two additional diagnoses of why he applies different ideas and approaches in each case: an *epistemology diagnosis* and an *accountability diagnosis*. We conclude with the possible implications of these alternative diagnoses for interdisciplinary education efforts.

Two Problems Asking "What Counts as a Good Approximation?"

As part of a larger study, we designed a set of problems to investigate how students might reason with similar mathematical content in different disciplinary problem contexts. This paper will include an in-depth discussion of one student's reasoning on two of these problems. The two problems, PENDULUM and ARCTAN, ask about approximations and infinite series in physics and calculus problem contexts, respectively:

> *PENDULUM (includes a figure of a pendulum with angle, length, and mass labeled): You have a pendulum made of a metal ball on a string. The string is 1 meter long and the metal ball has a mass of 1 kg. You might know that the approximation for the period of a pendulum*

for small oscillations is: $T = 2\pi\sqrt{\frac{l}{g}}$*, where T is the period of the pendulum, l is the length of the pendulum, and g is acceleration due to gravity (9.81 m/s²). This equation only holds for small angle oscillations of the pendulum. For larger angles, the period of a pendulum can be found with the following equation:* $T = 2\pi\sqrt{\frac{l}{g}}\left(1 + \frac{1}{16}{\theta_0}^2 + \frac{11}{3072}{\theta_0}^4 + \cdots\right)$*, where* θ_0 *is the angle of displacement of the pendulum from vertical in radians. You want to calculate the period of oscillation for this pendulum. How big can the angle of displacement of the pendulum be before the equation for small oscillations isn't a good approximation of the period?*

TAYLOR SERIES (note: since both problems relate to Taylor series and infinite series, for clarity, we refer to this problem in this paper as "ARCTAN"): The Taylor series about x = 0 for arctan(x) is: $arctan(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots$*. How big a value can x be before stopping at the second term is a bad approximation?*

Each problem contains two mathematical expressions, one an approximation of the other. In both problems, these two expressions are exactly equal when the relevant parameter is zero ($\theta_0 = 0$ in PENDULUM and $x = 0$ in ARCTAN). However, as these parameters increase, the two expressions become more different. The interviewee's task is to determine how large the parameters can get before the two expressions are so different that one is no longer a good approximation for the other. Although the specific mathematical expressions were not identical across both, these two problems deal with a similar issue of how to judge "what counts as a good approximation?" for an infinite series expression. Because the idea of what counts as a "good approximation" is not precisely defined here, tasks such as these reveal how interviewees might make these judgments differently in different disciplinary problem contexts.

Will's Reasoning on PENDULUM and ARCTAN

The interview with Will was conducted in the summer of 2011 by the first author. At the time, Will was a rising sophomore at a large, public east coast university and was planning on declaring a mechanical engineering major. In his freshman year, Will had completed two semesters of calculus (which covered infinite series and Taylor series) but had only taken one semester of physics (which did not cover pendulum oscillations). However, he had taken physics in high school and commented that he had seen the small angle approximation equation for the period of a pendulum before. The interview session lasted about 90 minutes. In the interview, Will reasoned about PENDULUM first (from 00:02:25 – 00:45:22), then ARCTAN (00:45:22 – 01:15:16), and then reflected on his work on the two problems (01:15:16 – 01:27:12).

The data collection and analysis were driven by two research questions: 1) "How do students approach these approximation problems similarly/differently in the different problem contexts?" and 2) "What factors plausibly support this similar/different reasoning?" The interview was semi-structured in that the interviewer was free to ask follow-up questions and probe deeper into unclear statements or interesting emergent topics. To address the research questions, we analyzed the entire video with corresponding transcript and written work to characterize Will's approaches to these problems as well as his thoughts and reflections on those solutions. Once we had these initial characterizations, we performed more careful, line-by-line analysis through key sections of the transcript, looking for confirmatory or disconfirmatory evidence of our characterizations.

Through this analysis, we argue that one difference between Will's approaches to the two problems is that he attempts to *make sense of and explore* possible analytical solutions on PENDULUM, while he instead attempts to *recall ideas* from calculus on ARCTAN. Then, we argue for three possible diagnoses of why such different approaches were taken: a *content knowledge diagnosis,* an *epistemology diagnosis*, and an *accountability diagnosis*. In conclusion we argue that these diagnoses have implications for how transfer of reasoning across the disciplines is supported.

Will's Approach to PENDULUM: *Making Sense and Exploring*

Will starts his work with PENDULUM by noting that he has not seen this problem before, and he reads the text carefully to make sense of what it is asking. Although unfamiliar to him, Will is able to make sense of the expressions and notes that being a good approximation means the two expressions for the period are equal to each other. He notices that the two expressions are the same except for the addition of $2\pi\sqrt{\frac{l}{g}}\left(\frac{1}{16}{\theta_0}^2 + \frac{11}{3072}{\theta_0}^4\right)$ in the series expression, and that each progressive term in the series gets smaller and smaller. He says that this additional term is added to "make up for the error that occurs…when the angle gets too big." He goes

on to explain that for small angles, the additional term is small, and so the two expressions will be approximately equal, but for larger angles, the additional term is large, so the two expressions will not be equal.

Will takes two approaches to trying to answer the problem: setting common-sense, physical bounds on the angle and attempting to set up and solve an equation that will produce a value for θ_0. In both approaches, he uses his physical understanding of the motion of the pendulum in order to evaluate the reasonableness of the answers. In setting an upper bound on the angle, he looks at the picture and declares that the upper bound is 90 degrees, because he feels that motion past that would no longer be that of a typical pendulum. While attempting to manipulate the equations to find values of θ_0 within the limits of a good approximation, one approach Will tries is to set the two expressions equal to each other to solve for θ_0. While the interviewer points out that $\theta_0 = 0$ would satisfy this mathematical relation, Will argues that is not a reasonable answer, because when the angle is zero the pendulum is not swinging at all. As Will says, "You can't really have a period when it's just sitting there not moving."

We label Will's approach here as *making sense and exploring*. Will starts by making sense of how two unfamiliar mathematical expressions are similar and different. Throughout the problem, he uses his understanding of the physical motion of a pendulum to make sense of the bounds on the angle. Without a sense of what the correct approach must be, Will tries to set up an equation and solve for θ_0, using his physical understanding to evaluate his results. Independent of his final answer here, this overall approach of *making sense and exploring* stands in contrast to his approach on ARCTAN.

Will's Approach to ARCTAN: *Recalling Formal Knowledge*

In contrast to making sense of an unfamiliar problem on PENDULUM, Will starts ARCTAN by pointing out that he does not remember the general method for finding the Taylor series of a function and by recognizing this problem as the kind he has done before in calculus class.

Throughout ARCTAN, Will fixates on certain pieces of knowledge he learned in his math classes that he believes are relevant and attempts to recall them. He mentions that arctangent has asymptotes, but he can't remember what they are. He cues in on the phrase "about $x = 0$" as being important and tries to remember what it means and how it is relevant for this problem. He also tries to remember common Taylor series expansions, such as those for sine and cosine, as well as the general formula for generating Taylor series.

As with PENDULUM, Will attempts to set bounds, on $\left(x - \frac{1}{3}x^3\right)$, for which the approximation is good. He vaguely remembers that some parameters have to be less than 1 in certain series convergence tests in calculus, and he uses this remembered fact to set the bounds as $0 < \left(x - \frac{1}{3}x^3\right) < 1$, although he recognizes that this is a guess and is not the "right way to do it." Later, the interviewer supplies two pieces of information at Will's request: the graph of arctangent with asymptotes labeled ($y = \pi/2$ and $y = -\pi/2$) and the general formula for Taylor series. Will uses the asymptotes of arctangent to change the bounds to $-\frac{\pi}{2} < \left(x - \frac{1}{3}x^3\right) < \frac{\pi}{2}$. He does not use the general Taylor series formula in his final answer.

Although he produces bounds for what counts as a good approximation in both PENDULUM and ARCTAN, his approaches are very different. Throughout his work on ARCTAN, rather than attempting to make sense of the relevant equations and ideas or trying to manipulate expressions to solve for x, Will seeks to *recall formal knowledge* from his calculus class. At the end of the interview, Will articulates his awareness of these two different approaches:

956 [01:17:00] W: [On ARCTAN] I was thinking from the top down in that, "What's a Taylor
957 series? What are the equations related to it? Uh, what's the information that I was once told
958 relating to this subject that I would need in order to find basic numbers, the number that I
959 need." So it was actually a complete opposite way of looking at it. [On PENDULUM] I
960 immediately read it and tried to, *pssh*, clear my mind and, alright, what am I looking for?
961 Basics. [On ARCTAN] I immediately read it and I was filled with, "What's Taylor series
962 again? How do you find the equation for approximations?"

Three Diagnoses of Why Will Does Not Take the Same Approaches to These Problems

How can we explain why Will attempts to *make sense and explore* on PENDULUM but tries to *recall formal knowledge* on ARCTAN? Here, we present three possible diagnoses.

1) A *Content Knowledge Diagnosis*

Attention to the content knowledge used and differences in specific problem features offers several specific diagnoses of why Will takes different approaches to these two problems, suggesting possible treatments for supporting alignment of Will's reasoning across these problems:

I) Although both problems involve infinite series, ARCTAN contains many explicit cues to Taylor series (i.e. explicit mention of "Taylor series" and the phrase "about $x = 0$"). On PENDULUM, Taylor series are never explicitly mentioned. If PENDULUM similarly contained these explicit cues to Taylor series ideas, Will may have sought to *recall formal knowledge* of Taylor series on PENDULUM as well.

II) Because $arctan(x) \approx \left(x - \frac{1}{3}x^3\right)$ includes transcendental and polynomial terms, it is not as easy to manipulate this equation to solve for x as it is in PENDULUM, where the analogous equation contains only polynomial terms. This difference in the mathematical structure of the two problems could explain a difference in Will's approaches. If $arctan(x)$ was changed to a different expression, such as $(1 - x)^{-1}$, Will may have also have sought to solve for x algebraically on ARCTAN.

III) In PENDULUM, the diagram of the pendulum cues the physical motion, providing a way to determine the bounds on the angle, whereas ARCTAN contains no such diagram. Although not cuing the same type of intuitive knowledge about physical motion, a graph of $y = arctan(x)$ given in the problem statement (rather than given later by the interviewer) could provide similar visual cues for *making sense and exploring* bounds.

Here, the intention is not to present all possible such diagnoses. Rather, the above examples are meant to illustrate what we take to be a *content knowledge diagnosis*: differences in reasoning on these problems stem from different content knowledge being cued by different problem features. We have argued that this type of underlying diagnosis is common in interdisciplinary course design efforts. For example, Dunn and Barbanel's (2000) co-taught multivariable calculus/electricity & magnetism course develops the mathematics and physics in parallel, with common language and notation. The hope is that content knowledge will be learned in a way that isn't compartmentalized to particular disciplinary contexts or problem cues, fostering transfer of that knowledge across disciplines.

2) An *Epistemology Diagnosis*

Providing an alternative to the content-focused perspectives on transfer, another perspective focuses on knowledge activation and transfer as related to individuals' *epistemologies* (or views on the nature of knowledge and knowing.) Hammer, Elby, Scherr, and Redish (2005) argue that what has been described as transfer is the activation of similar knowledge resources in various situations, and that this activation depends not only on content knowledge and problem features, but also on epistemological stances towards what kinds of knowledge are appropriate in different situations. For example, they describe two students taking different approaches to answering the question of how a scale reading of a person's weight changes in a moving elevator. Hammer et al. argue that one student's epistemological stance supported a formal computational approach to this problem, because she started by listing all the numerical quantities in the problem, apparently preparing for a calculation. Another student interrupted by asking, "Do we even need to do all that calculation?" and proceeded to describe her physical sense of how the elevator floor is falling away from the person as an elevator is accelerating downwards. This second student's interjection was a challenge not to the correctness of the numerical values listed, but to whether this *kind of approach* was necessary. Hammer et al. argue that this second kind of approach to the question stems from a different epistemological stance towards what kind of knowledge will be useful or productive here.

While this example might suggest that these students have fixed epistemologies that can only support one type of reasoning or activity, other research shows that individuals have access to many different epistemological stances, of which different ones can be cued in different moments to support different kinds of reasoning (Hammer & Elby, 2003; Rosenberg, Hammer, & Phelan, 2006). In this section, we argue that Will's different epistemological stances towards the two problems support his two different approaches.

"Logical Reasoning" Is Useful on PENDULUM

In evaluating his work on PENDULUM, Will recognizes that he would not get full credit for this explanation if it were an answer on a test, because he expects that problems on a physics test would require precise, learned methods, not just guesses at the bounds. In spite of this, he still thinks that he would get some partial credit, because the "logical reasoning" that he demonstrated is valued:

554 [00:42:26] W: A lot of teachers, if you put something down that sort of makes sense, if you
555 put something down, and you show logical thoughts and sort of show how you got to a semi-
556 close answer, they'll give you like a point or two and be like, "Alright, nice try, but not even

557 close." Um, and it's a big difference again between this stuff and, like, history. This kind of
558 stuff you can, even if you don't know it, you can use logic and you can, uh, make connections
559 and rationalize certain things and know that they're true just by looking at what you are given.
560 You don't need to know, I mean you've learned it over the years, but you don't need to know a
561 specific date or a specific event to answer a question. If they ask you, you know, "What's the
562 Battle of Hastings?" and you don't know anything about the Battle of Hastings, you're not,
563 you can't just be like, "There were swordsmen. They fight." You can't say that. But this one
564 you could say, "Alright I don't know what this equation really means, but I'm told, and I've
565 never seen it before, but I do know that T should equal T when the number of, when the, um,
566 degree of displacement of the oscillation is correct." And you can show stuff like that. And I
567 think if you show that you're willing to do that, they'll give you a little bit.

Here, Will's statements provide evidence of an epistemological stance that values "logical reasoning:" making sense of unfamiliar ideas, building connections, and providing evidence of rational thinking. Importantly, Will's statements are epistemological in that he is reflecting not only on his approach to PENDULUM, but also on *what kinds of approaches are appropriate* on this type of problem. This is reflected especially in contrast to the types of reasoning that are appropriate for historical questions, such as, "What's the Battle of Hastings?" which Will says cannot be reasoned through and requires specific prior knowledge.

Will's epistemological stance here is aligned with and plausibly supports his approach to PENDULUM. If he views "logical reasoning" as appropriate on this problem, then it is reasonable that Will would try to *make sense of and explore* this unfamiliar content rather than seek formal knowledge required to solve this problem. As a specific example, Will rejects $\theta_0 = 0$ as a possible answer to the problem, arguing that this answer must be wrong, not through a learned fact or procedure, but rather because of an intuitive physical interpretation that makes that answer violate common sense. In this way, Will's intuition-based evaluation of $\theta_0 = 0$ aligns with his epistemological stance that such "logical reasoning" is appropriate.

The "Pure Mathematical Reasoning" in ARCTAN Cannot Be Reasoned Through

While attempting to recall ideas from his calculus class, Will reflects on what is difficult about ARCTAN, revealing his epistemological stance that the reasoning required for this problem is "not normal reasoning:"

773 [01:03:14] W: Um, but yeah. It's 'cause I, I just don't remember those, uh, that information. I
774 don't think that, and I couldn't, this is, I hate, it's, that's why I hated these problems so much,
775 that I couldn't reason through them. I couldn't think, "Oh well, infinity, you know, this
776 happens." I just can't, it doesn't make any sense.
777 I: It's not like the previous problem, is what you're saying?
778 W: No, not at all. That's like concrete. That's like, ok, pendulum moving. I can see that.
779 That's why I hated Taylor series so much, is 'cause you can't see it. It's, it's just, it's like pure
780 mathematical reasoning that's, like, not normal reasoning. It's, you think about it a different
781 way. You can't just think as a person like, "Oh yeah, it's, pendulum swings to a certain point,
782 this happens." You have to think about it in terms of, like, infinity and what happens when
783 you go to infinity. That's like, I don't think that's like in, humans don't think like that
784 naturally, so you have to learn it.
785 I: Ok.
786 W: So I wouldn't be able to really, I, like, I couldn't get partial credit on this problem.

Here, Will describes an epistemological difference between appropriate approaches on PENDULUM and ARCTAN. The "pure mathematical reasoning" here is described in contrast to the "logical reasoning" on PENDULUM (lines 778 to 784). For Will, ARCTAN requires unnatural, formal, mathematical ways of thinking. Again, in comparison to PENDULUM, this means that Will would not expect to get any partial credit on this problem (line 786), because he cannot replicate the canonical reasoning from calculus class. In this way, Will's epistemological stance here is more similar to the one he takes towards a history question like "What's the Battle of Hastings?"

Will's epistemological stance towards ARCTAN aligns with *recalling formal knowledge*. It does not make sense to attempt "logical reasoning" as he does on PENDULUM if the mathematical content on ARCTAN is unnatural and cannot be reasoned through. Instead, the only productive approach would be to recall facts from calculus class. An epistemological stance that rejects "logical reasoning" as productive also helps explain why even after his initial failures to recall the relevant ideas he seeks, Will persists in recalling formal knowledge, instead of attempting the kind of *making sense and exploring* that he does on PENDULUM.

Importantly, the *epistemology diagnosis* suggests one reason why content or problem feature alignment alone may not be successful. Aligning only the content features of two problems as prescribed by the *content*

knowledge diagnosis may not be successful in fostering similar approaches to those problems if in working on those problems one takes different epistemological stances to what knowledge and approaches are useful.

3) An *Accountability Diagnosis*

Another approach to transfer focuses on the importance of motivational aspects: achievement goals, interest, and self-efficacy (Pugh & Bergin, 2006). In attending to these factors, we found how a sense of accountability to knowing and being able to recall relevant facts and approaches (and the related negative affect and feelings of low self-efficacy when he can't remember) can help explain Will's persistence in trying to *recall formal knowledge* on ARCTAN.

Will Does Not Feel Accountable for Knowing the Canonical Way of Solving PENDULUM

We argue that Will does not feel accountable for such canonical knowledge on PENDULUM, likely supporting his *making sense and exploring* on PENDULUM. Through the initial interaction with the interviewer, Will's focus shifts from knowing the "right way" to solve the problem to expressing his own understanding and seeking a solution that makes sense to him.

At the start of working on PENDULUM, Will reads the problem silently for about 80 seconds, pausing only to confirm with the interviewer that he is not required be familiar with this type of problem already. After reading, Will confirms his understanding of the two expressions with the interviewer. He then starts asking about what the problem wants him to do:

48 [00:04:26] W: Ok. [pause] So you don't want me to calc-, you just want me to answer the
49 question, right? I don't have to do anything?
50 I: Uh, yeah. Yeah, I mean, does it make sense, what they're asking?
51 W: Yeah, it makes sense. I mean, I can understand what they're trying to say.
52 I: Ok.
53 W: I haven't done oscillations, but I can understand the idea that they're trying to get through.

Here, the interviewer's redirection from what Will should be doing on the problem (lines 48 to 49) to Will's understanding of the problem (line 50) may help shift the focus to what Will understands. The interviewer's answers to Will's earlier clarification questions may also support an interpretation of the activity as the interviewer helping him make sense of the unfamiliar problem rather than testing what he already knows. Furthermore, the uninterrupted time spent silently reading could contribute to a feeling that Will has the opportunity to make sense of an unfamiliar problem, rather than immediately having to recall an answer. In line 53, he indicates that he has not "done oscillations," positioning himself as not accountable to this content.

After this excerpt, Will again spends one minute silently reading and thinking about the problem, after which he again starts to ask the interviewer if his understanding of the question is correct:

54 [00:05:47] W: Well, I would guess that these T's should be equal, right? If they're going to
55 be, they're both going to be accurate approximations.
56 I: Ok.
57 W: So at a certain point, when they're no longer roughly approximate is when your angle is
58 getting too big that this one [the small angle approximation] breaks down, and this is the one
59 [series equation] you have to use. Is that correct?
60 I: Um, well I mean, I mean, I just sort of want to figure out how you would think about it.
61 W: Ok.
62 I: So whether or not it's correct.
63 W: Ok, so you want me to explain how I would think about it.
64 I: Yeah.

After this except, Will's talk switches from asking confirmatory questions about the problem (as in lines 54 and 59) to stating his understanding of the question and how he would answer it. This shift in Will's talk suggests a corresponding shift in Will's interpretation of the interview situation, from an evaluation of his knowledge to one where his ideas are important and worth exploring.

This interpretation of the interview situation aligns with *making sense and exploring* on PENDULUM. If the purpose of the interview is to hear how Will would think about the problem, not holding him accountable to particular canonical ways of understanding and approaching this problem, then this would provide space and support for the *making sense and exploring* in which Will ultimately engages.

Will Feels Accountable for Recalling Ideas from Calculus Class on ARCTAN

Unlike the start of his reasoning on PENDULUM, Will does not initially approach ARCTAN by reading and making sense of the problem. Instead, he immediately indicates familiarity with this kind of problem:

602 [00:46:18] W: Um, I, don't you need to know how to formulate the Taylor series for, I don't
603 know, you don't, I guess. I don't even know how to formulate the Taylor series, so it's a good
604 thing they kind of gave it to me. Um, ok, I absolutely hate this stuff, but I have done this one
605 before. I have taken, like, calc 2, so I have done this before [Will reads the problem to
606 himself again].

Throughout the problem, Will signals that he feels accountable for remembering the relevant calculus content by indicating both his frustration at not being able to recall those facts as well his lower self-efficacy:

701 [00:58:28] W: And it's annoying, this one's really annoying, because I definitely have done
702 this or something like it, so I should know how to do this one. It's been in my mind before,
703 um, but I did get a 40 on this test, so, didn't know it that well. Uh, [laughs] let me think.

Unlike on PENDULUM, here the interviewer does not intervene to suggest to Will that he is not necessarily expecting him to recall facts from class, possibly because Will doesn't directly ask a question to the interviewer during ARCTAN. Instead, after Will attempts to recall information and work on the problem, the interviewer asks Will what information he would look up or ask someone, given the chance. This likely supports an interpretation by Will that prior knowledge is important, supporting his feelings of accountability. Similarly, receiving that knowledge later on may tacitly signal that this information was crucial.

At the end of Will's work on the problem, Will presents his answer, reflecting on how he feels about his answer:

880 [01:11:52] W: Um, but yeah, this one [ARCTAN] is more upsetting to me than the other one
881 [PENDULUM], because I did actually do these kinds of problems before. And like, I don't
882 really have, or see how to do these right, but that's how I would do it at this point, not
883 remembering much.

Will's sense of accountability towards knowing the canonical and correct methods and the associated feelings of frustration, displayed throughout his work on ARCTAN, provide another plausible support for why Will continues to seek to *recall formal knowledge*, even if that formal knowledge is not obviously useful for the problem. For example, throughout the problem he attempts to remember the general Taylor series formula, as well as the series expansions of sine and cosine, even though these are not necessary for solving the problem. A *content knowledge diagnosis* would suggest that cuing Taylor series simply cued the recall of other Taylor series ideas. An *accountability diagnosis* suggests that, on top of this, the goal of recalling facts that he learned in his calculus class in order to demonstrate competence and relieve frustration may support persistence in *recalling formal knowledge*. At the end of the interview, Will supports this interpretation when he explains why he tried to recall the general formula for Taylor series even though he did not use it in his answer: "I know I'd done it before. It was just frustrating me that I didn't remember the basic idea of it."

Importantly, this diagnosis differs from an *epistemology diagnosis*, because it looks beyond what kinds of knowledge and reasoning Will thinks the task requires. An *accountability diagnosis* also considers feelings of being accountable for that knowledge, such that not being able to remember those ideas leads to feelings of frustration and low self-efficacy, driving persistence in attempting to recall formal knowledge.

Implications of These Alternative Diagnoses for Interdisciplinary Education

Although research into transfer has opened up possibilities beyond the *content knowledge diagnosis*, it remains a common diagnosis in interdisciplinary efforts. In this paper, we raise the plausibility of other such diagnoses, namely an *epistemology diagnosis* and an *accountability diagnosis*, arguing that efforts that focus solely on alignment of content and problem features are incomplete. Here, we discuss some possible implications of each alternative diagnosis for interdisciplinary education efforts.

The *epistemology diagnosis* implies that instructors should also want students to view certain different disciplinary problems as requiring similar kinds of knowledge and approaches. Some interdisciplinary efforts have started incorporating the *epistemology diagnosis* into their course goals – for example, some recent physics courses aimed at life sciences students have the explicit epistemological goal of helping students see physics concepts and the associated mathematical reasoning as relevant and useful for understanding biology and have showed preliminary success in achieving these goals (Meredith & Bolker, 2012; Redish et al., in press). Although these courses do align particular topics to build connections between physics and biology content

knowledge, they simultaneously aim to help students view physics knowledge and ways of reasoning as productive for understanding biological problems, to support future transfer of physics into biology.

We have also shown that, for Will, feelings of accountability, with the associated low self-efficacy and negative affect, lead to brittle transfer of prior knowledge from his calculus class and persistence in unproductive approaches on ARCTAN. Although there have not been similar interdisciplinary reforms in response to an *accountability diagnosis*, some related research offers possible treatments. Elliott and Dweck (1988) mitigated the negative affect and low self-efficacy associated with "helplessness" by emphasizing the learning goals of an activity over goals of demonstrating competence. This led to both increased willingness to make public mistakes and improved performance over time. It is possible that deemphasizing performance goals could have similarly helped Will try different approaches on ARCTAN related to *making sense and exploring*. Interdisciplinary courses could aim to emphasize learning goals over performance goals, in order to counter low-risk/performance-oriented attitudes that may impede novel transfer across disciplinary courses (as well as to counter feelings of low self-efficacy, frustration, and helplessness that can impede learning more generally).

Our goal is to raise awareness of these multiple diagnoses of why transfer across disciplines may fail. Although we advocate for the inclusion of alternate diagnoses, we do not wish to downplay the importance of the *content knowledge diagnosis*, nor do we mean to suggest that these diagnoses are mutually exclusive. Surely, instruction that attends to *only* epistemology or *only* the low self-efficacy and negative affect related to feelings of accountability will be as incomplete as sole attention to content knowledge. Rather, enriching our understanding of the multiple reasons why transfer across disciplines can fail has the potential to lead to improvements in current interdisciplinary efforts. Additionally, it is possible that these multiple diagnoses can enrich our understanding of why *current* interdisciplinary efforts succeed, illuminating epistemological, affective, or motivational benefits of approaches derived purely from content considerations.

References

Al-Holou, N., Bilgutay, N. M., Corleto, C., Demel, J. T., Felder, R., Frair, K., … Wells, D. L. (1999). First-Year Integrated Curricula: Design Alternatives and Examples*. *Journal of Engineering Education*, *88*(4), 435–448.

Dunn, J. W., & Barbanel, J. (2000). One model for an integrated math/physics course focusing on electricity and magnetism and related calculus topics. *American Journal of Physics*, *68*(8), 749–757.

Elliott, E. S., & Dweck, C. S. (1988). Goals: an approach to motivation and achievement. *Journal of personality and social psychology*, *54*(1), 5–12.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*(3), 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Hammer, D., & Elby, A. (2003). Tapping Epistemological Resources for Learning Physics. *Journal of the Learning Sciences*, *12*(1), 53–90.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–120). Greenwich, CT: Information Age Publishing.

Meredith, D. C., & Bolker, J. A. (2012). Rounding off the cow: Challenges and successes in an interdisciplinary physics course for life science students. *American Journal of Physics*, *80*(10), 913–922.

Plomer, M., Jessen, K., Rangelov, G., & Meyer, M. (2010). Teaching physics in a physiologically meaningful manner. *Physical Review Special Topics - Physics Education Research*, *6*(2), 020116.

Pugh, K. J., & Bergin, D. A. (2006). Motivational Influences on Transfer. *Educational Psychologist*, *41*(3), 147–160.

Redish, E. F., Bauer, C., Carleton, K. L., Cooke, T. J., Cooper, M., Crouch, C. H., … Turpen, C. (in press). NEXUS/Physics: An interdisciplinary repurposing of physics for biologists. *American Journal of Physics*.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, *6*(3), 436–450.

Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, *115*(3), 381–400.

Rosenberg, S., Hammer, D., & Phelan, J. (2006). Multiple Epistemological Coherences in an Eighth-Grade Discussion of the Rock Cycle. *Journal of the Learning Sciences*, *15*(2), 261–292.

Acknowledgments

This work was supported by a sub-award under NSF-CCLI-0941191. We thank the TRUSE conference organizers for their support, Angie Little for her contributions to this research project, and Andy Elby, Ayush Gupta, Gina Quan, Joe Redish, and Vashti Sawtelle for thoughtful feedback on previous versions of this work.

Diving into Practice with Children and Undergraduates: A Cultural Historical Approach to Instantiating Making and Tinkering Activity in a Designed Learning Ecology

Lisa Hope Schwartz, University of Colorado Boulder, lisa.h.schwartz@colorado.edu
Daniela DiGiacomo, University of Colorado Boulder, daniela.digiacomo@colorado.edu
Kris D. Gutierrez, University of Colorado Boulder, kris.gutierrez@colorado.edu

Abstract: "Making and Tinkering" has become popular in informal education circles. The practice links science, technology, engineering and mathematics learning (STEM) to the do-it-yourself "maker" movement, where people of all ages "create and share things in both the digital and physical world" (Resnick & Rosenbaum, 2013). This paper examines a cultural historical approach to Making and Tinkering at two sites of El Pueblo Mágico (EPM), a social design experiment that joins university students, researchers and k-8 youth from predominately non-dominant communities together in joint activity for expansive and consequential learning (Gutierrez & Vossoughi, 2010). In the lineage of design experiments in the learning sciences our work addresses what Resnick and Rosenbaum discuss as the critical importance of designing contexts for tinkerability from a theory-based and iterative design approach that aims to both understand and change practice. With our work we seek to re-mediate normative STEM learning contexts for youth from non-dominant communities.

"Making and Tinkering" has become popular in informal education circles. The practice links science, technology, engineering and mathematics learning (STEM) to the do-it-yourself "maker" movement, where people of all ages "…create and share things, in both the physical world and the digital world" (Resnick & Rosenbaum, 2013, p. 163). This paper examines how undergraduates, children and researchers worked together to instantiate a cultural historical approach to making and tinkering at two instances of *El Pueblo Mágico* (EPM), a social design experiment that joins university students, researchers and k-8 youth from predominately non-dominant communities together in joint activity for expansive and consequential learning (Gutierrez & Vossoughi, 2010).

In the lineage of design experiments in the learning sciences our work addresses what Resnick and Rosenbaum (2013) discuss as the critical importance of "designing contexts for tinkerability (p. 179)", from a theory-based and iterative design approach, that aims to both understand *and* change practice and the context of learning (Gravemeijer & Cobb, 2006). Expanding the design experiment in what Gutierrez and Vossoughi (2010) discuss as a social design experiment, an overarching goal of our research is to address issues of equity by engaging theory in practice to circulate power in joint activity, so that participants work to collaboratively negotiate the process and objectives of their work. We articulate this process as re-mediation (Gutierrez, Hunter & Arzubiaga, 2009), or the redesign of what is normatively the social organization of STEM activity in academic spaces in order to expand and reconfigure the forms of mediation available for youth, in particular girls and youth from non-dominant communities, within STEM learning environments.

Following an emphasis in design research on situating a theoretically articulated approach in practice, in order to recursively generate new theoretical understandings and design principles, we first present our theoretical framework to Making and Tinkering (M & T) at *El Pueblo Mágico* as a hybrid model. Our approach engages commonalities across the Maker Movement as articulated by Resnick &Rosenbaum (2013), tenets of cultural historical theory that are integral to the designed learning ecology that is the context of our work (see Tables 1 and 2), and our research theorizing the "inventos" or inventive, everyday practices of Latino youth and families (Schwartz & Gutierrez, 2013). To examine our theoretical approach in practice at the sites of our research we investigate two interrelated questions**:** 1) How do undergraduates enact and conceptualize cultural historical concepts in M & T activity with children? and 2) How does our approach work to provide opportunities for participants to engage in robust STEM practices that involve children and adults in joint articulation of the objectives of activity? To explore these questions we present examples of what we view as effective appropriation of our theoretical approach as well as how less successful activity is reflective of the normative STEM education practices we are working to re-mediate. We argue that our cultural historical approach to M & T within our designed learning ecology presents a significant example for other researchers and practitioners who are working to create contexts for tinkerability that address issues of equity in STEM education.

A Cultural Historical Approach To Making and Tinkering

Our implementation of M & T at *El Pueblo Mágico* is based on tenets of cultural historical theory historically taken up in the social design experiments developed by Gutierrez (Gutierrez & Vossoughi, 2010). These activity systems involve undergraduates enrolled in university courses on child and adolescent development in joint

activity with children in an after school space. This multi-sited social design experiment is organized around what Gutierrez and Vossoughi (2010) call "equity-oriented and robust learning principles (p. 101) taken up "toward transformative ends through mutual relations of exchange (p. 101)". Of significance, our work seeks to re-mediate the functional system of science education for all students, and in particular for women and students from non-dominant communities. We do this through foregrounding the joint activity, playful inventiveness, and human ingenuity we see as common threads of theory and practice across social design experiments, maker spaces, and the activity of members of non-dominant communities.

An overarching goal of both EPM and M & T articulated by Resnick and Rosenbaum (2013) is to re-mediate normative school practice. Resnick and Rosenbaum explain that rather than a focus on planning that is often pervasive in academic contexts and that adheres to formal rules and recipes for activity, tinkering is messier, more improvisational and more open ended, and is actually more like how science is actually practiced by scientists. What they refer to as "bottom and up" forms of participation align with cultural historical activity theory's (Engestrom, 1987) emphasis on horizontal and vertical development that underlies EPM. In other words, both approaches seek to leverage participation in a range of activities for deepening, broadening and connecting interests, and for forming sustained inquiries through multidimensional movement. The biggest shared criticism of M & T and the EPM design experiment from teachers and novice teachers respectively is that activity seems unorganized and unstructured. We concur with both Resnick and Rosenbaum that this view is misguided because "true tinkerers (p. 167)" know how to move tinkering into focused activity. It is precisely this capacity that we are working to foment in teachers and children at EPM. Rather than teaching lesson planning and top down classroom control, we emphasize "diving into practice" with children and a process of mediated praxis (Gutierrez & Vossoughi, 2010) whereby novice teachers reflect on activity and visualize new possibilities for joint activity with children.

Table 1. A Cultural Historical Approach to M & T at *El Pueblo Mágico*:

El Pueblo /4411 Learning Ecology	**Making and Tinkering**	**Practices in Homes**
Play and the imaginary situation as forming zopeds (Vygotsky, 1978)	Play; Experimental, iterative style of engagement	Play is the main activity videotaped in 14 homes
Stone & Gutierrez (2007) - Joint problem articulation - Serial mediation e.g. continual reassessment and re-directing of object-oriented activity - Just enough assistance Gutierrez & Vossoughi (2010) - mediated praxis	- Continual goal reassessment - Continual exploration of new paths and imagining new possibilities - Immediate feedback	Schwartz & Gutierrez, 2013 - Directives - Assistance for "leveling up" - Tool sharing - Idea sharing - Assisted turn-taking
Expansive learning (Engestrom, 1987) - horizontal / vertical movement - growing together everyday and scientific concepts (Vygotsky, 1978)	- Fluid experimentation: easy to dive in, connect and extend - Process over product - Open exploration - Improvisation/adaptation/iteration	Schwartz & Gutierrez, 2013 *Inventos*: crafting new rules of engagement with digital media to engage tight circumstances
Community of Leaners (Rogoff, 1994) - Distributed expertise among intergenerational ensembles - Learning as taking on new roles and responsibilities in joint activity (Vygotsky, 1978) - Cultural mediation	- Engagement with people and materials	- Distributed expertise, fluidity of roles between expert and novice - Assistance for novice players - Shared / negotiated access to tools within families
(Gutierrez, 1999) - Hybridity and heterogeneity	- Diverse examples, divergent thinking	- Hybrid language practices

Table 1 above outlines our approach to M & T within our designed learning ecology. Table rows are organized to show complementary concepts across the three approaches and activity systems. Column one outlines the main tenets of the cultural historical approach articulated in the university course. Column two, derived from Resnick and Rosenbaum's seminal paper, shows how our approach connects to and draws new emphases from M & T. Column three shows how we pull from our research on digital media and learning in homes (Schwartz & Gutierrez, 2013). Overall, each of these contexts and frameworks highlights play, joint mediated activity with people and things, hybridity, joint problem articulation, distributed expertise and fluid experimentation.

Putting the Design into Practice: Context and Methods of Data Collection and Analysis

In the sections that follow we focus on the work of undergraduates, called *amigos* (friends) at EPM, and children in the two instantiations of our approach to M & T within the social design experiment of *El Pueblo Mágico*. Table 2 and Figure 2 and 3 show information about each context. In both sites of EPM, members of the research team served as designers and facilitators of activity. To differentiate these two contexts we use 1) EPM1 for the Spring 2013 instantiation with undergraduates and predominately Latino children from non-dominant communities in grades 2-5 at Posada elementary school, and 2) EPM2 for the summer 2013 site with undergraduates and middle school students from privileged and non-dominant communities at Smiley Middle School. At this site, Schwartz conceptualized the M & T activity and coordinated the research team. Fundamental to the EPM social design experiment is providing space for reflection and dialogue between the undergraduate and the instructor. In EPM1 the cognitive ethnography (CE) that engages students in dialogic reflection on activity is the key space for mediated praxis. At EPM2 this occurred primarily through interaction, blog posts and course papers. Consequently, our primary data sources are CEs for EPM1 and videotaped data, course papers and blog posts for EPM2.

In order to address our questions regarding how undergraduates enacted and conceptualized M & T we first conducted an overall analysis of CEs for EPM1 and videotaped data for EPM2 (see Figures 1 and 2). For EPM1, a subset of CEs was identified through text searches for M & T and key concepts taught in the course (e.g. mediation). Out of 132 CEs by 26 students, where students averaged 5-6 CEs/EPM visits, 86 (65%) of CEs by 24 (92%) of students discussed M & T, and of this subset, 76 (88%) of CEs by 23 (96%) of students used key theoretical concepts. Subsequently, we selected the work of a subset of students (n=8) that demonstrated a range of effectiveness in putting theory into practice for further analysis. We chose 26 CEs written by five undergraduates from EPM1 and the work of three ensembles at EPM2. For EPM2, initial analysis and coding was a team effort between Schwartz and our undergraduate research opportunities (UROP) team of physics and education students. We assigned a ranking scheme and notation for undergraduates' questioning strategies and indicators of group collaboration. After selecting representative cases we coded data with a focus on 1) types of strategic assistance, 2) distributed expertise and roles emerging in activity, and 3) use of key theoretical concepts for mediated praxis.

Table 2. EPM1 and EPM2, Site, Participants, Key Activities and Data Sources

Sites	Participants	Key Activities / Terms	Data Sources
EPM1 - 5507 child development course - EPM at Posada Elementary School over one semester	- Undergraduates in 5507 (one time a week over 2.5 months) - 5507 instructors - EPM staff (doctoral students) - Children in grades k2-5 (~85) - Researchers / M & T support	Adventure Guides: - Zoom Zoom (cars) - Scribble machines - Squishy circuits - AgentCubes / Sheets - World Maker	- 86 Cognitive Ethnographies (CEs) by 24 students - X CEs by 4 students for close analysis - Video data; field notes
EPM2 - 5508 adolescent development course - EPM at Smiley Middle School / CU Boulder, 6 days over 2 weeks	- Undergraduates in 5508 - Children grades 6 and 7 (18) - Researchers, 5508 instructors, and M & T and EPM2 designers and facilitators - UROP student designers and researchers	- Solar Cars, solar theramin - Produce circuits - LED/squishy circuits - Sewn circuits - Minecraft circuits	- Videotaped data from 6 days of EPM2, ~1.5 hours per day - Utilized closely three 8-18 minute long clips - Field notes - Student papers, blogs

Strategies and Concepts for Learning and Becoming in Practice

Figures 1 and 2 below show the big picture regarding the theoretical concepts and forms of assistance utilized by undergraduates as well as instructors. On closer examination of the work of 8 students in EPM1 and EPM2, our analysis revealed that successful strategies for re-organizing the roles and responsibilities, or what we term distributing expertise in activity and jointly articulating problems, involved strategic questioning, in particular what we designate as "design questions". We found that that successful design questions allowed for children to "dive into practice" and extend their inquiries by focusing children's attention and imagination on particular

aspects of the task, re-voicing children's nascent ideas to the group, and giving options for potential directions for activity that included children's choices and interests. Undergraduates reflected on their activity and the role of play, motivation and engagement primarily through the concepts of mediation (and related forms, e.g. serial mediation, re-mediation), zoped, and community of learners. The examples shared here show how these students also used terms resonant of M & T, such as "dove into" and "trial and error". In both instances of EPM, moving away from a planning and recipe approach and taking up fluid experimentation where children could immediately engage with tinkering and were subsequently supported in reflecting on and extending their practices was acknowledged by participants as a critical re-mediation of accustomed school practice. Activity discussed below demonstrates how the common tensions felt by teachers to have all the answers and to serve as the sole authority were re-mediated by the process of jointly articulating the objectives of activity with children through distributing responsibility for thinking, imagining, teaching and learning to undergraduates and students.

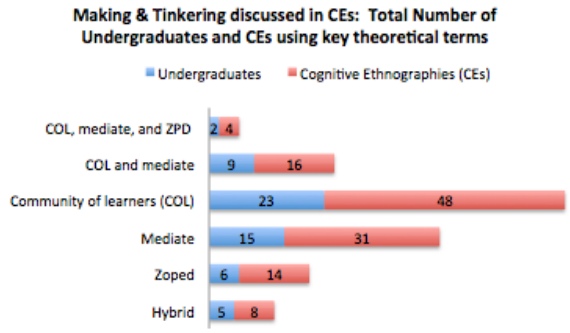


Figure 1. Concepts used by undergraduates regarding M & T at EPM1

Undergraduate assistance: Questioning, direct assistance, modeling, side-by-side, doing nothing, sharing ideas	
1: good	12 instances identified
2: adequate	11
3: poor	9
Peer support: Modeling, turn-taking, sharing ideas, etc.	
1: good	10 instances identified
2: adequate	9
3: poor	5

Figure 2. Ranking of assistance strategies in 32 instances of interaction at EPM2

We present an example from Ann Smith's CE to show how activity created new participation pathways for children. Smith documented her group's creation of "squishy circuits" with playdoh and LED lights. She explained how she turned thinking over to the students, through idea sharing and questions eliciting their thoughts about design. Smith related how this provided the space for 2nd grader Cecilia to take on a new voice and role in activity:

> I asked them if they all remembered how to make the Squishy Circuits and Flor and Cecilia said they did, but Michael told me he had never made them before and asked me how to make it. (OC: This is where I thought that making the other kids the in group the expert instead of me would be a better way of getting the instructions across). Cecilia, who usually doesn't talk much, piped right up and started explaining to Michael how the Playdoh had to be on top of the insulating dough and the Playdoh couldn't touch other Playdoh or it wouldn't work. Then she said that the battery wires had to be touching the Playdoh, but not the insulating dough and that the light had to be plugged into those same Playdoh pieces. (OC *…it was a nice change to hear her talk more than I had ever heard her talk before. Cecilia also acted as the mediator in this process between the instructions and Matthew understanding how to make the circuit.*). Michael looked like he kind of understood what Cecilia had said, but tried to pretend that he understood everything because he dove right into making a mermaid.

This example demonstrates how Smith privileged distributed expertise. Significantly, Cecilia a Latina girl and second grader, who Smith related was usually extremely reticent, became the expert teaching an older boy. She gave Michael what Stone & Gutierrez (2007) call "just enough assistance" for him to dive into making his circuit. Smith recounted "Their interactions also showed Vygotsky's ZPD. Matthew was not able to make his lights turn on until Cecilia turned his light the other way. This simple act of assistance showed me that Cecilia understood how the circuits worked and was able to help Matthew come to that same understanding." Cecilia continued to provide assistance to Michael until he eventually completed a circuit on his own. The interaction shows how consciously distributing expertise to students and allowing them to take on new responsibilities supported fluid experimentation and the creation of zopeds that engaged students' potential development. Importantly, activity in Smith's group supported a young Latina girl, a member of two groups (women and Latinos) underrepresented in many scientific fields in taking on the role of an expert. The following example also demonstrates expanded possibilities for normative gender roles with 3rd grader Maria taking a leading role.

In the next example, Schwartz supported an undergraduate, Suz Miller and 3rd grader Maria, on strategies for joint problem articulation. Maria wanted to create squishy circuits but was resisting group work

and getting started. Schwartz, through privileging joint activity among a wider range of participants, assisted their team with how they might collaborate with a group that was creating a movie. She suggested that Maria might contribute to the movie by helping to fabricate set items the group wanted with squishy circuits materials. In her CE, Miller described the learning opportunities that were opened up by widening the frame of possibility for collaboration:

> …Maria used her experiences with the scribbling machine to communicate its function to the group. No one else had done the tinkering activity so they were all novices making her the expert. The children's roles swapped while filming as Maria had a very minimal understanding of that project. Operating in a diverse group promoted the members zone's of proximal development as they acquired the opportunity to apply knowledge across many activities. Problem solving through group trial and error produced unique solutions as the ensemble members exchanged ideas and learned together. The opportunity to revise activities further enhanced critical thinking and the transfer of knowledge. Our problem solving process resembled a reflective collaborative learning model as the undergrads initiated communication and the children expanded on topics/ ideas.

Despite Maria's initial reluctance to join the new group, Miller related that the merger was extremely successful primarily through the cross-pollination of ideas, and distribution of expertise and roles among participants. Miller's work with the children also utilized what we term design questions to mediate joint problem articulation:

> "Oh, so you think we should lay the propeller flat like this instead of attaching it upright like a wing? What do the rest of you think?" I "What feature of the machine do you think needs to change in order to make it fly?" [OC: using open ended questions I guided the children's thought processes and re-structured my questions when they did not seem to grasp what I originally presented]. "It needs to have four spinning things not two, like a helicopter," Maria suggested. "So you think we need more propellers, and Isaiah thinks the propeller needs to be attached differently. Should we try these theories out and see if they work?" [OC: Maria used her understanding of flying objects to construct an analogy that helped her articulate her hypothesis to the group].

Miller's open ended questions about the design of the "flying boots" for the film assisted children in connecting their thinking to prior experiences and to features of the design needed for their current objectives. She specifically asked children what they thought and modeled taking up others' divergent thinking as resources for activity. Miller also used scientific language and practices to suggest to the students to test out their ideas with continued tinkering. Overall, her strategic questions distributed expertise to the children and expanded their activity. The open-ended but specific function of the questions that Miller used to turn decision making over to the children are what we define as features of "design questions". In her CE, Miller discussed her question-asking strategy with the concept of mediated serial assistance (Stone & Gutierrez, 2007), a process where the facilitator helps to organize interaction so children jointly determine the sub-tasks and direction of activity. She wrote "mediated-serial assistance appeared far more often in my group this week…As we worked through the flying machine issue I promoted critical thinking by posing "open-ended" questions to the group. … as the children responded I acknowledged their ideas, reflected on them, and expanded on the question in new ways".

In each of these examples, a focus on design and the imaginary situation engaged children in fluid experimentation whereby they could jump into activity, but also pull back and reflect on the direction of their goals. By helping children imagine possibility, undergraduates used play to form a zoped that engaged students' potential understandings. Additionally, undergraduates' discourse moves provided immediate feedback that did not restrict children's imagination and helped push them into new perspectives and practices.

Problems with Planning

Within M & T and EPM, play can be hard work, but there is an element of spontaneity and experimentation that enlists participation. In CE1 Ruth Penn wrote her group attempted to begin the adventure guide for Zoom Zoom (car creation). She related that Andres and Jose easily shared ideas and drew complex sketches, but that they quickly lost interest in planning and went to play computer games. She surmised that if they had immediately worked on building the car, they might have stuck with the activity. In CE3 her group was undecided on a project. Penn related how an EPM staff person intervened to lecture them that they had to create a plan before they began "worldmaker". The dialogue shut down the boys' enthusiasm, with Andres's lack of voice indexed by the action of putting a piece of tape over his mouth. This situation, where an EPM staff-person presented an approach contrary to the one espoused in this article, is also indicative of the gaps that often arise between

university and k12 schools, the hurdles for putting theory into practice, and how easily novice teachers will refer to a perceived authority, even if the interaction does not align with theory they are learning in their university courses.

In CE4 Penn wrote that Andres exclaimed, "I can't believe I am having fun!" Her final CEs show a move away from planning and the introduction of new tools to immediately tinker with. In these CEs Penn's discourse evolved into more strategic questioning and idea sharing. In CE5, Penn explained her role in Andres and Miguel's design of an experiment with Mentos candy and soda.

> … the new goal became how to make the soda explode higher, while putting in as many Mentos as possible before it exploded. Once we went inside the library, my assistance turned to become a mediator. I mediated the discussion we had with the experiment. When discussing what we would do for the future, and why the soda exploded, I asked open-ended questions to encourage critical thinking. I didn't want to give them the answer, but at the same time, if I didn't ask them to think about what they would change in the future, I'm not sure if they would have volunteered the information. So, I acted as the lever to get them to answer these important questions (CE6).

The progression of activity in Penn's CEs shows how when children are restricted from tinkering and articulating he goals of activity, they lose interest and motivation, but when Penn allowed for the youth to jump into activity first, she was able to position herself, through questioning and framing the boys inquiry, to serve as the "lever" for expanding their thinking. Unsurprisingly, when Andres was having fun, he was engaged and extended his abilities as he created and reflected on his world building in AgentCubes and his "candy chemistry" experiments.

Hands, Control and Distributed Expertise

Enjoyment and learning also went hand in hand at EPM2. On the last day Schwartz asked children to relate what they had learned and what they enjoyed. Children shared they liked designing and testing their cars, learning about circuits, building and competing with others in popular digital game Minecraft, and working with undergraduate amigos. Many students expressed that what they enjoyed and learned were the same. This section discusses joint activity at EPM2 among middle school aged youth and undergraduates in an adolescent development course. EPM2 focused specifically on M & T activities as exemplar practices for learning theory within the university course and at EPM. Our discussion highlights best practices and problematic exchanges in instantiating our approach to M & T. Table 3 below outlines the activity of the three ensembles we selected for further analysis. In each group children from non-dominant communities immigrant heritage were paired with Anglo youth and undergraduates, one of whom was also Latino. These ensembles present a range of forms of assistance, from primarily top-down, to a mix of directives and guided design questions, to more open-ended assistance.

First we focus on the movement of hands on materials in the different ensembles. For solar car construction there were tensions around the control of materials. In groups two and three the instructors utilized more heavy-handed "next-step" design questions—questions that veered towards known answers and lacked a distribution of agency to the children for deciding the direction of subtasks (Stone & Gutierrez, 2007). The most problematic interaction occurred in group three. It began with Amber constructing the solar car while the three boys looked on:

> Instructor Bill: Why are your hands all over it?
> Amber: I was trying to put the wheel on
> Instructor Bill: Why are *you* trying to put the wheel on?
> Amber: Because they can't do it
>
> Instructor Bill: (playful tone) What do you mean they're capable 8th graders with working hands.

This exchange re-mediated Amber's top-down approach and the boys took control of the car materials. But, shortly thereafter activity moved back to Amber and Instructor Jim who took the materials to demonstrate when he saw the boys having trouble. Watching the adults work together indexed their authority for Jorge. He exclaimed, "you guys are so smart!". Eventually thinking and acting was turned back over to the youth when Marnie initiated two design questions modeled after Bill's and with his support negotiated a shared placement of the car's motor. Activity in ensemble two was less problematic but also relied on directives and "next-step" design questions. As seen in row two, the children had double the amount of time on turn with the materials as Tamara. Still, Manuel's time with the materials overshadowed Yolanda's turn-taking, even as he oriented more than Yolanda to Tamara's approval. A positive aspect of this group was their use of hybrid language practices.

During the most interactive sequence between the children they utilized Spanish to discuss their shared decision-making.

Table 3. The Activity of Three Ensembles at EPM2

	Group 1	Group 2 (3 people)	Group 3
1. Participants	Undergrad: Marnie 3 boys: Merza, Tarik and Tom	Undergrad: Tamara Manuel and Yolanda	Undergrad: Amber 3 boys: Edgar, Bob, and Joe
2. Materials	Produce Circuits: Multiple sets	Solar Cars: One set	Solar Cars: one set
3. "Hands On": Times on turn with materials, and implying manipulation of materials	Undergraduate, Tarik and Sam: each have their hands on their own materials Merza: observes (related that Merza was able to explain the whole process later)	Undergraduate: 8 Manuel: 14 Yolanda: 4 turns (touched materials 15x) Girl & Undergrad: 4 Boy and girl: 1	Undergraduate: 10 Instructor Jim: 5 Jorge: 8 Bob: 5 Joe: 0 Adult total =15 Children =13 Among Instructors: 1 Undergrad & Boys: 2
4. Primary Undergraduate Discourse Strategies	- Models her own thinking (8) - Questions boys to explain their thinking (18) - Suggests boys view each other's work (10) - Refers to prior experiences (8)	Explicit Directives (7) "Next step" design questions, e.g. ""how will the wheel turn?, "where does this go?" (7)	Explicit Directives (2) Yes / No questions (3) IRE (2) Next step design questions (9) Design questions (3)
5. Role of Course Instructor(s)	Bill: Strategies and ideas for participants thinking, models discourse for Marnie	No instructor present in interaction	Bill: Re-mediates top-down approach, design questions Jim: Models, questions

In group one the movement of people and expertise was much more fluid and open. Merza and Tarik moved constantly, and Marnie followed suit, in order to engage with them. Marnie referred to the children's prior experience, during the summer program, and more broadly in their lives to assist the creation of produce circuits:

Marnie: When it didn't work last time with the play do, what did we do to the light?
Tarik: We switched it.
Marnie: We switched it. Do you wanna try to switch that and see what happens?
Tarik: So… (Mumbles. Sticks LED into playdoh, pauses). This is what we did with the playdoh
when the light didn't work. (Pulls LED out, turns it around, sticks it back in)
Marnie: hmmm (points, touches LED) What could be wrong? I wanna have you trouble shoot it.
Tom: Maybe the bulb burned out?
Marnie: The bulbs burned out? Okay, lets try a different bulb.
Tom: (Puts a new bulb in the circuit, it works).
Tarik: Ah I knew it worked!
Marnie: Awesome you just made another circuit. Congrats! How can you use that to extend it?
Tarik: You can connect another battery

Marnie's discourse encouraged Tarik to trouble shoot. The tone of interaction remained playful despite initial lack of success. When the LED did not light up, Tarik was not blamed, rather it was recognized that the issue may be with the materials and not the user. This interaction literally ignited Tarik's confidence. He took up Marnie's invitation to extend his tinkering, exclaiming "Ooo!" when Marnie obtained more batteries for him to use. Marnie's learning was also supported during this activity. Akin to the situation with Amber, Instructor Bill offered examples of questions to elicit students thinking and agency that Marnie used when the group struggled.

Tinkering With Our Design

At EPM the complex layers of mediation from children, undergraduates, and instructors present us with rich resources to support continued tinkering with our cultural historical approach to M & T. We saw positive outcomes regarding the re-mediation of participation in STEM activities for youth from non-dominant communities when undergraduates allowed for children to dive into activity, and when their strategic questioning and assistance distributed thinking to children. While these may seem like very simple strategies, unfortunately we also saw in our own sites how often this is not the approach adults take with children. Undergraduates who were successful in re-mediating STEM activity invoked theories of serial mediation and

joint problem articulation and related how their strategies worked to distribute expertise to children so that the youth could form zopeds with their peers.

The kind of multidimensional movement and distribution of expertise that we place as the central affordance of our approach to M & T necessitates examining contradictions that arise in activity. While we see in homes that children's *inventos* foment the creation of new rules for distributing participation when materials are scarce, at home and at EPM we see that it is often far too easy for the older person or the male to take the role of authority and owner of the materials and thereby claim the expertise to dictate the objective of activity. Though we discussed here some examples where young girls took the lead in teaching others, there are also more gender normative exchanges, including the predominance of Manuel's hands on the car materials. Continuing to examine and design for broadening normative gender roles and rules is critical to the evolution of our approach. Aligned with this issue is continued thinking on how to design new rules, roles and artifacts for distributing materials and expertise so that joint problem articulation is embedded in interaction.

We find with M & T what might be "good" strategies in one context can easily remove agency from participants in other interactions. Design questions can serve as supports for students thinking and distribute agency to youth through "just enough" assistance (Stone & Gutierrez, 2007) to prompt their continued inquiry. Yet, through slight changes in wording or through tone or gesticulation, questions around design can work in the more lockstep fashion of next step assistance. We plan to continue retrospective analysis of design questions and discourse with more of our video data and CEs. Through the present analysis we see that many of the contradictions that arise with design questions are linked to issues of confidence. Children and undergraduates worry about being right, and need to be supported in taking risks and not giving up authority to those who normatively wield power. However, in this model, adults *do* need to take the role of facilitator who can widen the frame of possibility and model thinking through sharing ideas and ways to approach open-ended problem solving. A continued tension, also related to this role, is the amount of expertise the undergraduates need in order to extend children's thinking, and how much they should be involved in the activity of tinkering to create their own or shared products with children.

In our next iteration of the design experiment we plan to set up mobile M & T materials centers coupled with undergraduate "connectors" who model generative design questions and idea sharing. The centers will provide spaces for children and undergraduates to jump into activity together. Through presenting an array of materials we will organize the space for horizontal and vertical expansion across activities and ways of knowing. Additionally, we have begun an extension study where we visit children's homes to work with families on M &T activities. With this endeavor we seek to continue to engage families' *inventos* as a source for M & T, and to expand participation through actively valuing and enlisting a wide range of practices for STEM learning. To conclude, we offer a fundamental cultural historical concept that informs this social design experiment: "change in the individual involves change in the social situation itself" (Engeström, 2008 in Gutierrez & Vossoughi, 2010, p. 101).

References

Engeström, Y. (1987). Learning by Expanding. Helsinki, Finland: Orienta-konsultit.

Engeström, Y. (2008). From teams to knots: Activity theoretical studies of collaboration and learning at work. New York: Cam- bridge University Press. Engeström,

Gravemeijer, K. & Cobb, P. (2006). Design research from a learning design perspective. In: Van den Akker, J., Gravemeijer, K, McKenney, S. & Nieveen, N. (Eds). (2006). *Educational design research*. London: Routledge, 17-51.

Gutierrez, K., Hunter, J. & Arzubiaga, A. (2009) Re-mediating the University. *Pedagogies: An International Journal* 4 (1-23).

Gutierrez, K. D., & Vossoughi, S. (2009). Lifting Off the Ground to Return Anew: Mediated Praxis, Transformative Learning, and Social Design Experiments. *Journal of Teacher Education*, *61*(1-2), 100–117.

Resnick, M., & Rosenbaum, E. (2013). Designing for Tinkerability. In Honey, M., & Kanter, D. (eds.), *Design, Make, Play: Growing the Next Generation of STEM Innovators*, pp. 163-181. Routledge.

Schwartz, L. & Gutierrez K. (2013, March). Turn-Taking and *Inventos*: Examining The Everyday Lives of Latino Families and Designing Learning Ecologies with Youth and Under-graduates. Paper presented at the Digital Media and Learning Conference, Chicago, IL.

Stone, L.D. & Gutiérrez, K.D. (2007). Problem articulation and the processes of assistance: An activity theoretic view of mediation in game play. *International Journal of Educational Research 46*. 43-56.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes (*pp. 79-91). Cambridge: Harvard University Press.

Facilitating Design Research by Mapping Design Research Trajectories

Guanzhong Ma, Jan van Aalst
The University of Hong Kong, Pokfulam, Hong Kong SAR, China
Email: gzma@connect.hku.hk, vanaalst@hku.hk

Abstract: Design research is an emerging paradigm in the study of learning. It is far from a mature methodology, as it faces a variety of difficulties, one of which is the characterization of the research process. We mapped the design trajectory proposed by Sandoval to capture the complex process of design research. The design trajectory map develops Sandoval's conjecture mapping in two ways, with the capacity to capture movement along the trajectory. We apply the proposed mapping to a well-known design study that described the design history of an inquiry-based learning project. The limitations of the proposed mapping approach are discussed.

Introduction

Sandoval (2004, 2014) proposed *embodied conjecture* and developed the technique of *conjecture mapping*. An embodied conjecture specifies how theoretically derived conjectures about how learning occurs can be reified as a concrete design in a learning environment. The technique of conjecture mapping involves the visual representation of conjectures about how design elements are predicted to bring about the desired learning outcomes. Continuous evaluation of the enactment of the design results in the refinement of the conjectures, which serves to develop theories of learning. The evolution of conjectures about learning theories during the various iterations in a design-based study constitutes the design research trajectory. In other words, conjecture maps that represent the evolution of conjectures document the design research trajectory.

The goal of this paper is to further develop conjecture mapping in two ways. First, design research requires not only the *proposing* of conjectures about how a design will work, but also the testing of these conjectures. Sandoval's conjecture mapping (Sandoval, 2014) acknowledged the need for testing conjectures, but to date has not represented this in a conjecture map. A sequence of conjecture maps along a timeline represents the design research trajectory, but does not evaluate the result of the conjectured mechanisms. Nevertheless, the evolution of conjectures results from testing them. We propose that the extent to which the predicted relationships between design elements and outcomes are observed during testing should also be mapped to provide a more comprehensive documentation of the design research trajectory, and to present how the conjectured learning theories are empirically supported. We developed a technique with such a capability by mapping the evolving degree to which the conjecture relationship is achieved. We hope that this attempt addresses Sandoval's (2014) concern that "the current formulation of conjecture maps does not easily capture movement along a research trajectory" (p. 34). Second, in the mapping, we addressed the challenges observed in a design's implementation by taking into account the role of conjectures in informing potential refinements of the design, thus further enriching the conceptualization of the design research trajectory.

In the next section we describe the components of the design research trajectory to be mapped. Guidelines for the mapping are then provided and applied to a seminal report on design research (Edelson, Gordin, & Pea, 1999), which examined an early attempt to use scientific visualization technologies in high-school earth science courses. Although this is not a *recent* design study, it remains valuable because of the number of iterations involved and the amount of detail reported regarding the design decisions. What can be gained from the mapping is then discussed.

Mapping the Design Research Trajectory

Documentation of the Design Research Trajectory

In the design research trajectory, the design is implemented in a targeted context: researchers monitor how the design works in that context and make decisions to revise, modify, and refine the design to achieve the desired outcomes. In this trajectory, knowledge relevant to the design, context, and learning is developed. Documentation of the design research trajectory is suggested by researchers to show the "trace of the decision making" (Joseph, Bryk, Bransford, Gomez, & the Information Infrastructure Project, 2003). This practice is intended not only to guide communication and coordination across the contexts of different studies, but also to allow research audiences to evaluate the credibility of design decisions and the quality of lessons learned from the research (Collins, Joseph, & Bielaczyc, 2004; Joseph, 2004). In addition, documentation of the evolution of design ideas and the underlying rationale may inform the broader community's understanding of the research

process, which is essential for distinguishing educational design research from educational design (McKenney, Reeves, & Herrington, 2012).

Documentation of the design research trajectory needs to be based on a thorough understanding of the features of design research. Design research involves continuous evaluations (Edelson, 2002) and iterative refinements (Collins et al., 2004), so a sequence of evolving maps that represent the evaluation results of the conjectured components could be used to document the design research trajectory. Each map represents the evaluation result of a particular phase in the research. The sequence of evolving maps aims to document the longitudinal process of design, enactment, and refinement, which responds to Sandoval's (2014) call for capturing movement along the research trajectory. This documentation of the longitudinal process is required for design research, because recent instructional research has illustrated the fluid and non-linear nature of educational design processes.

As both conjecture development and conjecture testing are involved in design research, mapping the conjectures and the evaluation results is needed to represent the design research trajectory. Sandoval's (2004, 2014) work addressed the mapping of conjectures, whereas our focus is on the mapping of evaluation results. Combining the maps of Sandoval (2014) with those proposed here for the same research would produce a more comprehensive representation of the design research trajectory.

Components in the Trajectory to Be Mapped

Previous attempts to document the design research trajectory have addressed different features of the design research process (Barab et al., 2002; Collins et al., 2004). These attempts did not adequately link the processes of enactment to outcomes, which may impede the development of theories that explain why a design succeeds or fails in supporting learning (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Design-Based Research Collective, 2003). Sandoval (2004) proposed the idea of embodied conjecture, and argued that developing a conjecture map would facilitate the refinement of the conjectures to develop learning theories that explain how a design brings about learning outcomes. The conjecture map consists of design elements, intermediate outcomes, and intervention outcomes. Intervention outcomes refer to the "typical sort of outcomes that psychologists look for, like whether students learn what they are intended to learn" and intermediate outcomes refer to "observable patterns of behavior predicted by a model of how an embodied conjecture functions to support learning" (p. 215). Observation of intermediate outcomes in the enactment of a design is the basis of the achievement of intervention outcomes. If an intervention outcome is not observed, then this suggests that some intermediate outcome may be absent in the enactment.

Predicted relationships between design elements and intermediate outcomes, and between intermediate outcomes and intervention outcomes, are mapped to specify how the conjectures are embodied within the learning environment to support learning. Sandoval (2014) further developed the technique to capture the longitudinal progress of design and enactment from his earlier work (Sandoval, 2004), in which the components of the conjecture map were represented statically rather than being updated as the enactment proceeded.

Sandoval's conjecture map articulates the links among particular interactions between design elements and mediating processes (intermediate outcomes), and between mediating processes and intervention outcomes. That is, the conjecture map represents a hypothesized set of causal links specifying all of the intermediate outcomes that each particular design element is predicted to bring about, and all of the intervention outcomes that each particular intermediate outcome is predicted to bring about. Moreover, these causal links also specify all of the predicted causes of each intermediate outcome or intervention outcome. As the design is evaluated indirectly through its predicted impacts on the intermediate and intervention outcomes, the specification of the causal links in the conjecture map enables researchers to retrospectively identify the weaknesses of a particular design element when a predicted outcome is not observed after implementation. This may facilitate the iterative refinement of the design.

The components addressed in the conjecture map articulate some salient features of the design research trajectory. However, considering the features of design research, we argue that *challenges* relating to preconditions and constraints in terms of realizing desired learning should also be mapped. Design research can be characterized as a process in which a designed product or artifact is placed in the context of its use to obtain feedback for further refinement. In the process of enactment, we should not expect that the designed intervention will promote the intended learning behavior and processes without encountering challenges. McKenney et al. (2012) suggest that in design research, constraints such as not being able to change standards or assessment methods and limited participant time should be taken into account when developing the design. In addition, as already mentioned, design research typically introduces an innovative learning environment with the expectation that students will change how they learn. This change may be even greater if students learn in a technology-enhanced learning environment that differs substantially from those that they experienced previously. These challenges may, as a result, block the achievement of intermediate and intervention outcomes. For the documentation of the design research trajectory to better inform design research, the challenges should not be overlooked when developing and evaluating conjectures, and thus should be incorporated in the map.

In addition, the design research trajectory involves both developing and evaluating conjectures. The extent to which the design produces the desired learning outcomes, as represented by the conjectures, should be mapped in the interest of refining the conjectures and the design. McKenney et al. (2012) reviewed existing models and frameworks depicting the design research process (Jonassen, Cernusca, & Ionas, 2007; Reinking & Bradley, 2008) and found that many consider problem characterization, design or development, and evaluation or empirical testing, as necessary phases. This seems to require documentation of the design research trajectory—a series of characterizations of each of the phases. It is not sufficient only to characterize how the designed learning environment is expected to support learning; it is also necessary to depict the results of the evaluation of its enactment. Doing this informs subsequent refinement of the design and the problem characterization, and results in updated conjectures about how the design would work in the context of use. Sandoval's conjecture maps specify how such conjectures evolve over the course of design research. However, although informed by evaluation of the enactment, the sequence of conjecture maps does not depict the results of that evaluation. Thus, maps representing the results of continuous evaluation complement conjecture maps in documenting the design research trajectory.

In considering the mapping presented here, we found it useful to draw on Sandoval's (2004, 2014) conjecture maps. This paper uses a series of maps to depict how the components and the links in the design research trajectory evolve. We highlight some of the features here to introduce the maps. In the next section, we demonstrate how the maps were developed, using a design research example for illustration.

Mapping the Design Research Trajectory: An Example

This section describes an example of mapping the design research trajectory. We briefly explain the selection of a design research study for analysis, and summarize that study, and then describe the resulting design research trajectory maps.

Construction of the Mapping

The construction of design research trajectory mapping involves identifying the components and determining whether and how the predicted relationships between the components as given by the conjecture are empirically supported.

Several features of the mapping should be highlighted. *A sequence of maps* is used to represent the longitudinal process by which the design and insights about it evolve. Each map in the sequence represents the results of evaluation of the conjectures in the corresponding phase of the design research. We found it useful to draw on Sandoval's (2004) conjecture maps when considering which components needed to be mapped; thus, each map shows design elements, challenges, intermediate outcomes, and intervention outcomes. The design elements in a map result either from the initial design and the underlying conjecture, or are added as a refinement to the original design. Intervention outcomes represent what students are intended to learn, such as conceptual understanding of subject matter (Sandoval, 2004). We also added the *challenges* existing between the design elements and the intermediate outcomes to show how the context of use may impede the achievement of intermediate outcomes. Challenges represent preconditions or barriers that make the targeted learning difficult to realize. They can be identified through teacher-implementers' comments on the instruction, researchers' reflection on the design, and students' reports on their learning and assessment results. Finally, we used different styles of arrows to indicate the extent to which a predicted relationship was achieved, representing the results of the evaluation of the conjectures. For example, arrows were used to show whether a particular design element contributed to address a particular challenge and the extent to which overcoming a particular challenge contributed to a particular intermediate outcome. In summary, the sequence of maps serves to document the design research trajectory with a focus on continuous evaluation during each phase of the design research for development and validation of the conjectures.

There is a lack of consensus on the rubric for determining to what extent a predicted relationship between two components is empirically supported based on the information in the research report. To exemplify the mapping, we propose the following guidelines. These guidelines are not intended to be universally applicable to design research or to fit best with the mapping for this research example. It is also noted that although the research example we analyzed here is one of only a few in design research that details the research trajectory, the information reported was still insufficient to determine the extent to which each link in the diagram is empirically supported.

To evaluate the extent to which the conjecture is confirmed, we developed four guidelines. (1) The extent to which a predicted relationship is achieved is determined by how this relationship is empirically supported. (2) Three levels are used to indicate the extent to which a predicted relationship is confirmed—"fully supported," "partially supported," and "unsupported"; these are represented in the map by a solid arrow, a dashed arrow, and a dash-and-dot arrow, respectively. (3) A relationship is designated as fully supported if it is considered to have no significant problems, as partially supported if it could be further improved or an improvement solution is in progress, and as unsupported if it is not observed. (4) The maps are read from left to

right, such that the achievement of a left-hand component assumes achievement of the successive component to its right. In other words, if a left-hand component is not fully achieved and the component that is predicted to be achieved, this is indicated by an unsupported relationship between them unless specific information is provided.

The Study under Analysis

The literature contains a substantial number of studies purporting to be design research. However, many papers reported single case studies that do not involve iterative improvement of the design and the underlying conjectures (Krajcik et al., 1998; VanSledright, 2002). In other papers such refinements were reported, but these refinements were not always based on formative assessments of the design in its iterations (Bell, Hoadley, & Linn, 2004; Kolodner et al., 2003; Zhang, Scardamalia, Reeve, & Messina, 2009). For example, Zhang, Scardamalia, Reeve and Messina (2009) reported a post-hoc analysis of three successive social arrangements in students' on-line work with Knowledge Forum, implemented by the same teacher and with the same curriculum in three successive school years. Although this study provides important evidence of the effectiveness of each design, design changes were based on the teacher's satisfaction with each design, with empirical analyses being carried out later. Furthermore, most of the aforementioned papers in fact report on *research programs* consisting of multiple studies and publications. Findings may be reported in many articles without necessarily clearly discussing the relationship of the study to previous studies from the same research program, which makes the research trajectory difficult to understand. A succinct way in which to depict these connections thus seems necessary.

Edelson, Gordin and Pea (1999) did describe the relationships between their evidence, obtained from formative assessments and refinement-oriented decisions, although they did not report the details of the formative evaluations. We thus selected this study to illustrate the mapping procedures. Edelson et al. (1999) presented a design history of software and a curriculum that aimed to use scientific visualization technologies to support students' inquiry-based learning in geoscience. The researchers aimed to design a learning environment in which students could conduct authentic scientific inquiry as scientists in a laboratory. The topic of inquiry in the study was climatology (e.g., weather and global warming). Students worked on this topic by analyzing large collections of authentic quantitative data that were provided. The data were displayed in the form of rectangular arrays that used colors to represent ranges of numerical or categorical values, designated as scientific visualization. All of the data and tools for investigation were integrated in a designed software environment. Students were expected to investigate specific problems in this environment, through which process they achieved objectives such as general inquiry abilities, specific investigation skills, and the understanding of science concepts through participation. As the implementation of the design proceeded, challenges were recognized. The researchers iteratively refined the software and curriculum activities to address these challenges and to promote the desired outcomes. As a result, four versions of the design were developed and implemented, each constituting one phase of the research. Formative evaluation was conducted for each version. Accordingly, we may obtain evidence on the extent to which the design brought about the desired outcomes as conjectured in each of the four phases of the research.

Developing the Map

We highlight some of the features in the first phase of the research example and illustrate how the design trajectory can be mapped. In describing the mapping, we first introduce the components of the map, comprising design elements, challenges, and intermediate and intervention outcomes, and the predicted relationships between them. Thereafter, we map how these predicted relationships were achieved, as indicated by evaluating the design. As Sandoval (2014) devoted a paper to discussing the mapping of components in the design research trajectory, we provide more details here about mapping the evaluation of the predicted relationships.

The first version of the design was the Climate Visualizer (Gordin, Polman, & Pea, 1994). Figure 1 shows the map for the first phase of the research, read from left to right. Here, we focus on representing how the design would lead to the desired outcomes as conjectured. *Datasets* for inquiry in a data library related to weather information, a topic assumed to interest students, to address the challenge of *motivation* to engage in inquiry. A *supportive user-interface* was developed to address the challenge of *accessibility* of investigation techniques. Students were expected to *understand the techniques available* to them well, which would bring about the acquisition of *specific investigation skills* and an *understanding of the science content*. In addition, assuming that the students would have difficulty in planning, organizing, and coordinating the inquiry process, the researchers developed the Collaboratory Notebook to overcome the challenge of *managing extended activities* in inquiry-based learning. This environment was a *structured hypermedia environment*, in which students could plan and record their investigations, coordinate work efforts among collaborative teams, and receive feedback from teachers and mentors. *Technological effort* (e.g., software architecture) was made to address the *practical constraints* in the real classroom, such as available resources and fixed schedules. It was expected that by facilitating this, students could do better in *recording the process and products of the investigation*, which contributed to the improvement of their *general inquiry abilities*.

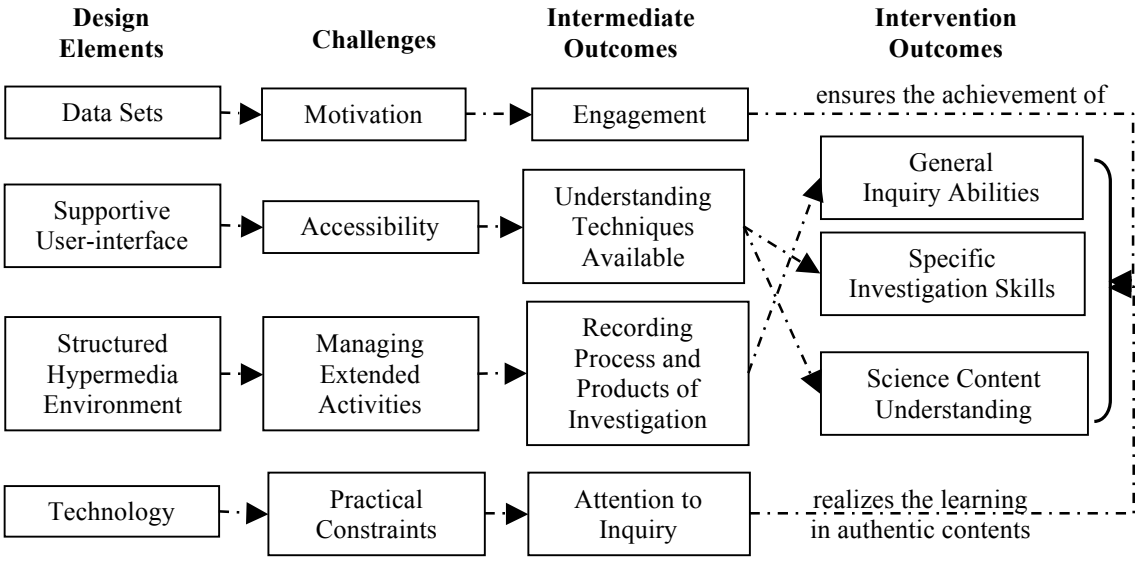


Figure1. Design trajectory map for the first version of the design in the research example

The first version of the design was reported to have "failed to provide teachers and students with an understanding of the full range of visualization techniques" (p. 410), and "failed to address student motivation" (p. 411). In addition, the software ran so slowly that it failed to draw students' attention to the inquiry itself, which was seen as "a practical failure to meet the constraints of the classroom" (p. 412). In other words, the link between the "datasets" as a "design element" and the "motivation" challenge, and the link between the "accessibility" as a "challenge" and the "understanding techniques available" as an "intermediate outcome," were both unsupported (dash-and-dot arrows).

The other relationships in the map were considered unsupported (dash-and-dot arrows) if their left-hand components were not achieved, unless specific information was provided. In general, insufficient information was provided in the original text of the research example, so that the map had to be developed literally according to the text. For example, information about whether the structured hypermedia environment addressed the challenge of managing extended activities was not presented in the paper. We thus assumed that in the first version of the design, all of the links between the components were not supported empirically, as no sufficiently specific information was provided.

Evolution of the Design Research Trajectory

The mapping of the first phase of the research shows that the predicted relationships between the design elements and outcomes were not well supported empirically. To address the potential problems observed, Edelson et al. (1999) made refinements to their conjectures, and developed second and third versions, the Radiation-Budget Visualizer and the Greenhouse Effect Visualizer (Gordin, Edelson, & Pea, 1995). Several design elements were revised, such as changing datasets to focus on more interesting topics, and the development of curricular activities.

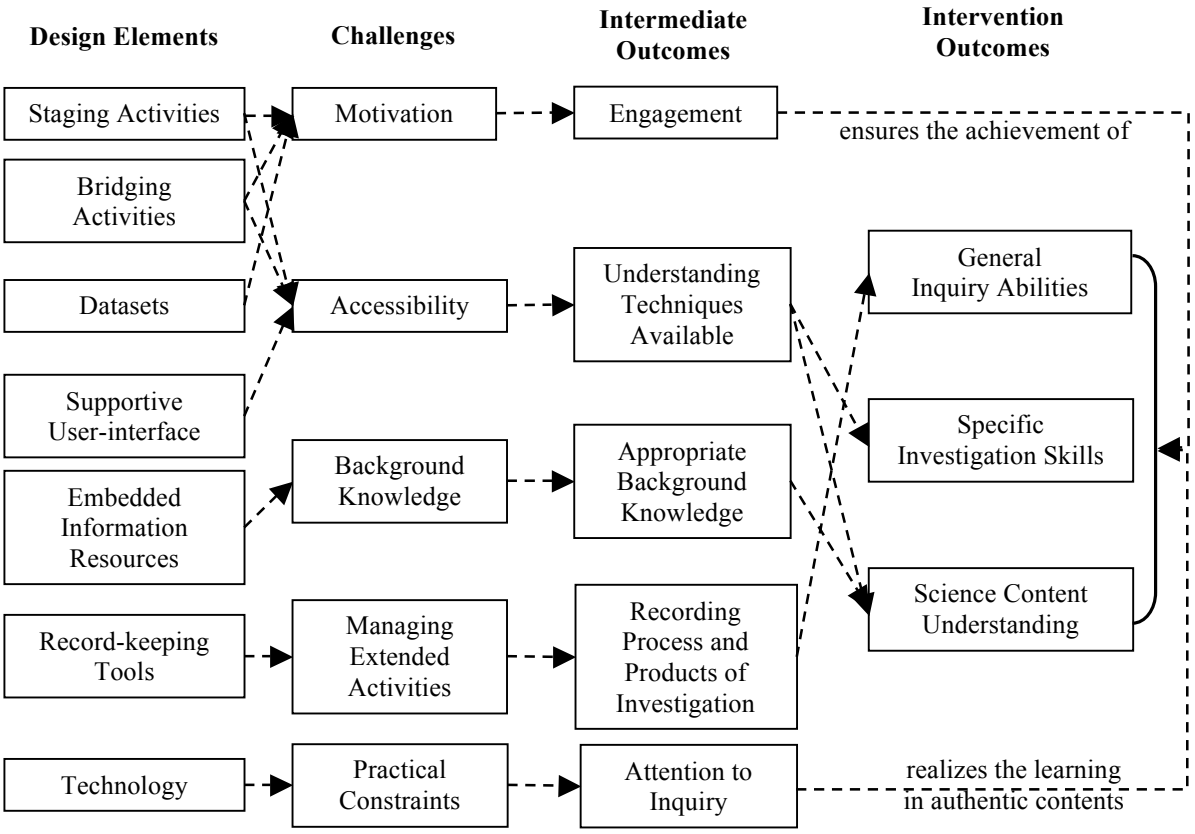


Figure 2. Design trajectory map for the fourth version of the design in the research example

To illustrate how the *mapping* represents the evolution of the design research trajectory, we describe the mapping for the fourth phase of the study. The fourth version of the design, WorldWatcher, further addressed the challenges through a series of attempts. The map of this phase is shown in Figure 2. We highlight the following features. First, the researchers expected that broadening the datasets, improving the user-interface, and introducing staging activities into the curriculum to motivate students to investigate a wider range of topics of interest could address the challenge of motivation. For example, the database was broadened to include datasets that supported a much broader range of investigations, which provided meaningful problems for students to inquire into. With this challenge addressed it was expected that students would engage in inquiry, which assumed the achievement of the desired learning objectives. Second, the design attempts in the previous phases of the study aimed at addressing the challenge of accessibility included improvement of the user-interface to help interpret the visualizations and the development of staging and bridging curriculum activities to allow the learning and practice of investigation techniques. For example, the staging activities provided a context within which students practiced using the techniques. The supportive interface, along with the specific operations in the software environment, provided students with access to powerful investigation techniques such as performing quantitative analyses of the data and viewing data in other visual representations. These attempts were expected to equip students with specific investigation skills, through the performance of which they would come to understand science content knowledge. Third, the researchers expected that the embedding of information resources and staging activities would alleviate the challenge of background knowledge for inquiry. Equipping students with appropriate background knowledge would help them to understand the science content knowledge. Fourth, the researchers provided various forms of record-keeping tools in the software and expected these to alleviate the challenge of managing extended activities. This made it possible to record the process and products of the investigations, which contributed to the learning of general inquiry abilities. Finally, the researchers aimed to improve the performance of the software to address the challenge of practical constraints, which would help realize learning in authentic contexts.

As reported, the challenges encountered in the enactment, such as motivation, accessibility, background knowledge, managing extended activities, and practical constraints, were not revealed as being as serious as those encountered in the enactment of the previous design. However, problems still existed, as seen from the researchers' observations and reports on the enactment. For example, students were observed to have difficulties with retaining sight of the overall inquiry context as they became involved in individual activities, which required reestablishing the motivating context. The information reported in the text was not sufficient to

determine how each of the links between components was empirically supported, since "little formal evaluation has been conducted" (p. 438) on the uses of this version. Given that issues raised by classroom experiences existed and that this design was being revised, it seems more appropriate to consider the links in the map as partially supported (dashed arrows).

Looking at the maps of the first and the fourth phases as a sequence, we can observe a progression in the design and the conjectures. First, the conceptualization of the design research trajectory evolved. The components and relationships mapped conceptualize the design trajectory by showing the salient features of each design. As shown in the sequence of maps, components were added, modified, replaced, and refined, and the relationships between the components were revised and validated. As a result, these researchers obtained a better understanding of the interactions between the design and the context of its use. Second, the theoretical knowledge of why the designs succeeded or failed in promoting learning outcomes also evolved. The maps were intended to represent the extent to which the predicted relationships between the design and the desired outcomes were supported. The processes of enactment were thus linked to the outcomes, which is expected to permit theoretical knowledge to be developed concerning why the design succeeds or fails in supporting learning. As shown, each map represents the extent to which the conjectured theoretical knowledge was empirically supported in each phase. Thus the sequence of maps presents the way in which this theoretical knowledge developed in clarity and was validated. In summary, creating a sequence of maps, each representing the reality in one phase of the research, was helpful for documenting both the conceptualization of the design research trajectory and the development of theoretical knowledge.

Discussion

Sandoval (2004, 2014) proposed the idea of embodied conjecture and developed the technique of conjecture mapping. His conjecture maps document the evolution of conjectures about how design elements are predicted to work together to support learning over the course of design research. The components in the sequence of conjecture maps characterize the salient features of the design research trajectory that links the design elements to the processes of enactment and to the desired learning outcomes. Conjecture maps thus facilitate systematic design research and the development of learning theories. This paper draws primarily on the technique of conjecture mapping, but makes some revisions intended to enhance the role of the mapping in facilitating design research.

First, we mapped the extent to which the predicted relationships between the design and the desired outcomes were supported. This supplements Sandoval's conjecture maps, as not only were conjectures about how the design would work to support learning mapped, but also how these conjectures were empirically supported. This revision is expected to provide a more comprehensive representation of the design research trajectory. Second, we incorporated a challenge component between the design elements and the intermediate outcomes. This component reflects the role of constraints in the context of use in impeding the achievement of intermediate outcomes. We suggest that mapping the challenges would be helpful for informing the design and evaluation stages in design research. We agree with Sandoval (2014) that the conjecture map represents an argument. A sequence of conjecture maps visualizes the hypothesized causal processes that link a design to its outcomes. These causal links may enable researchers to "look forward," by reading from left to right, along the pathways from design elements to desired outcomes. In the maps in the present work, the representation of the extent to which such causal links were validated informed how success or failure in observing some outcome could be traced back to a particular design element, when read from right to left. This paper contributes to the development of the technique of conjecture mapping by including the capability of capturing the movement of the design along a research trajectory, thus documenting a more comprehensive design research trajectory.

We acknowledge that the work reported here needs further improvement. For example, the issue of determining the level of support for the conjectures remains to be addressed. Sandoval (2004) distinguished between design-oriented conjectures and theoretical conjectures. It is reasonable to assume that different definitions and different standards for level of support are required for these two types of conjecture, given the difference in nature between them. The rubrics for determining the level of support as presented also need further clarification. In addition, as the design trajectory maps were constructed based on post-hoc analysis of published work, it remains to be seen how such maps could be useful in guiding new iterative design research. Research on the practice of applying such mapping in new design research projects is necessary for the evaluation and refinement of the technique.

References

Barab, S. A., Baek, E.-O., Schatz, S., Moore, J., Sluder, K., & Scheckler, R. (2002). Illuminating the braids of change in a web-supported community: A design experiment by any other name. In *Proceedings of the American Educational Research Association Conference* (Vol. 2). New Orleans, LA.

Bell, P., Hoadley, C. M., & Linn, M. C. (2004). Design-based research in education. *Internet Environments for Science Education*, 73–85.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher*, *32*, 9–13.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *Journal of the Learning Sciences*, *13*, 15–42.

Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 31(1), 5–8.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*, *11*, 105–121.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, *8*, 391–450.

Gordin, D. N., Edelson, D. C., & Pea, R. D. (1995). The Greenhouse Effect Visualizer: A tool for the science classroom. *Preprints of the Fourth American Meteorological Society Education Symposium*, Dallas, Texas.

Gordin, D. N., Polman, J. L., & Pea, R. D. (1994). The Climate Visualizer: Sense-making through scientific visualization. *Journal of Science Education and Technology*, *3*, 203–226.

Jonassen, D., Cernusca, D., & Ionas, G. (2007). Constructivism and instructional design: The emergence of the learning sciences and design research. *Trends and Issues in Instructional Design and Technology*, *2*, 45–52.

Joseph, D. (2004). The practice of design-based research: Uncovering the interplay between design, research, and the real-world Context. *Educational Psychologist*, *39*, 235–242.

Joseph, D., Bryk, A., Bransford, J., Gomez, L., & the Information Infrastructure Project (2003). *Information infrastructures*. Presented at the Annual Meeting of the European Association for Research on Learning and Instruction. Padua, Italy.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., … & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning by Design into practice. *Journal of the Learning Sciences*, *12*, 495–547.

Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J., & Soloway, E. (1998). Inquiry in Project-Based Science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, *7*, 313–350.

McKenney, S. E., Reeves, T. C., & Herrington, J. (2012). *Conducting Educational Design Research*. London: Routledge.

Reinking, D., & Bradley, B. A. (2008). *On formative and design experiments: Approaches to language and literacy research*. New York: Teachers College Press.

Sandoval, W. A. (2004). Developing learning theory by refining conjectures embodied in educational designs. *Educational Psychologist*, *39*, 213–223.

Sandoval, W. A. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*, 18–36.

The Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 31(1), 5–8.

VanSledright, B. A. (2002). Fifth graders investigating history in the classroom: Results from a researcher-practitioner design experiment. *Elementary School Journal*, *103*(2), 131–160.

Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for Collective Cognitive Responsibility in Knowledge-Building Communities. *Journal of the Learning Sciences*, *18*, 7–44.

Connected Gaming: Towards Integrating Instructionist and Constructionist Approaches in K-12 Serious Gaming

Yasmin B. Kafai, University of Pennsylvania, 3700 Walnut St, Philadelphia, PA 19106, kafai@upenn.edu
Quinn Burke, College of Charleston, 86 Wentworth Street, Charleston SC, 29425 burkeqq@cofc.edu

Abstract: In this conceptual paper, we argue that K-12 serious gaming should focus on connected gaming, which is signaled by a move from a predominantly instructionist focus on having students *play* educational games for learning to an increasingly constructionist approach that has students *make* their own games for learning. Constructionist activities have always been part of the larger gaming ecology but have traditionally received far less attention than their instructionist counterparts. We argue that constructionist gaming approaches promote not only meaningful content and collaboration but also creative and critical skills in the context of coding. We propose that future discussions of serious gaming ought to be inclusive of constructionist approaches to better realize the full potential of gaming as a means to genuinely connect children to technology and to each other and how this potential for more meaningful connectivity can address the persistent access and diversity issues long facing gaming cultures.

Introduction

> Every educator must have felt some envy watching children playing video games: If only that energy could be mobilized in the service of learning something that the educator values. But the envy can take very different forms. Instructionists show their orientation by concretizing the wish as a desire for games that will teach math or spelling or geography or whatever. The Constructionist mind is revealed when the wish leads to imagining children making the games instead of just playing them. Rather than wanting games to instruct children they yearn to see children construct games.
>
> *Seymour Papert (p. ii, 1995)*

Papert's assessment was remarkably prescient of developments that would follow just a decade later when the serious games movement started. Despite video games having become a multi-billion dollar business equaling—if not surpassing—the movie industry, gaming is still regularly dismissed by some educators as a waste of time, or even worse, an instigator of stereotypes and violence. In response to such critics, some theorists (Gentile & Gentile, 2008) have wondered that if video games are, in fact, such effective inculcators of stereotypes and violence, why can't their influence be harnessed for good and serve as powerful tools to support children's learning? This was the question that Gee (2003) answered when examining what video games could teach us about learning and literacy, arguing that many good educational principles—36 in total—could be found in the design and play of video games.

The launch of serious gaming realized the yearning for instructionist games, those games that are designed to teach academic content to students. Hundreds, if not thousands, of educational games and simulations have been designed to support learning in various domains (Shaffer, 2007; Squire, 2011). Accompanying these efforts were the launch of several conferences and journals, the funding of numerous research initiatives, and even the placement of a senior policy advisor on games and gaming in the White House. Following a report by the National Research Council (2011), a flurry of reviews have recently come out examining the learning benefits of instructionist games. The verdict reached by these meta-analyses is decidedly mixed: while one meta-analysis found significant impact (Wouters, van Nimwegen, von Oostendorp & van der Spek, 2013), others were more hesitant in their assessment of impact (e.g., Girard, Ecalle, & Magant, 2012; Vogel, Vogel, Cannon-Bowers, Bowers, Muse & Wright, 2013), while still others were downright dismissive of the motivation and cognitive benefits claimed by serious gaming (e.g., Young, Slota, Cutter, Jalette, Mullin, Lai, Simenoni, Tran & Yukhymenko, 2012).

There has however been one notable absence in all of these reviews: the inclusion of constructionist gaming approaches—namely those approaches in which games are designed by students rather than professionals for their learning benefits (Kafai, 1995; 2006). And this absence is surprising given the successes of constructionist gaming for not only learning programming but also academic content and other skills (see Hayes & Games, 2008). It is worth reflecting for a moment on what might have caused this omission. The first and most obvious reason stems back to what Papert aptly described as the instructionist desire of having a finished, downloadable, teaching product—namely, the game itself—as the party responsible (rather than the instructor) for teaching the child. A second and less inimical reason may be that constructionist gaming has been

less popular simply because educators have viewed the endeavor as far too technical given its association with learning programming. And last, a third reason may be that until recently, the gaming industry did not want players to engage in any design or modification of the games they produced for the marketplace. However, whatever the reasons—educational, technical, or cultural—the situation is clearly changing.

We are currently witnessing a paradigmatic shift toward constructionist gaming that is propelled by several developments, including the initiative to promote computational thinking (Grover & Pea, 2013), a need to broaden participation in computing (NRC, 2011), and the emergence of a DIY culture (Lankshear & Knobel, 2010). But the central impetus for a shift comes from the industry itself. After all, some of the most popular games on the market today include level and character modding as a central feature (El Nasr & Smith, 2006) and encourage such modding as part of game play until the next version becomes available. This element of constructionism is not limited to gameplay itself. A closer examination of gaming cultures reveals that many rich learning activities happen in the context of what Gee (2003) refers to as "metagaming" in which play extends beyond the game and includes participating in online discussion forums (Steinkuehler & Duncan, 2008) and even accessing cheat sites (Kafai & Fields, 2013) to help players more effectively navigate the game. In the community of many instructional game designers, we also observe a recent shift to include game making platforms and activities (Klopfer & Haas, 2012). Perhaps though the clearest indicator that constructionist gaming has arrived is signaled by the remarkable popularity of Minecraft (Duncan, 2011), a virtual sandbox whose tremendous popularity has garnered over 12 million paying designers and even served as the topic of a recent South Park episode.

In this conceptual paper, we articulate a long overdue framework for constructionist gaming to outline its learning benefits in terms of coding, creativity, collaboration, and criticality. Through this framework, we make a case for connected gaming, an approach to serious gaming that includes both instructionist and constructionist perspectives such that playing and making games are no longer treated as two separate activities but overlapping, mutually informing processes for learning. Connected gaming, as we argue, sees learning to play and make games as part of a larger gaming ecology in which the traditional roles of game player and game maker are no longer treated as distinct entities. In the following sections, we first conceptualize how the four components, or 4 C's of constructionist gaming, manifest themselves in the game-making process, and then outline a sample scenario for connected gaming, before addressing some of the key challenges in making the process more accessible to non-programmers and traditionally underrepresented groups. Our goal is to focus on how the learning sciences of serious gaming can be more inclusive and informative for children by giving young players a greater hand in the design and production of video games.

The "4C's" of Constructionist Gaming

Our approach to constructionist gaming builds on prior efforts to understand how and what children learn in the process of designing and making digital media through computer programming (Kafai & Burke, 2014). While current developments situate game making in several different fields such as new media literacies (Gee, 2010), system-based thinking (Salen, 2007) and critical engagement with media (Buckingham & Burn, 2007; Pelletier, 2009), we draw on the broader notion of participatory culture informed by Jenkins and colleagues' (2006) work. We identify four different dimensions of participatory competencies—*c*oding, *c*reative, *c*ollaborative, and *c*ritical—that are all relevant to constructionist gaming (Kafai & Peppler, 2011) and underpin the nature of connected gaming.

Coding

Coding is the most distinctive skill to be learned, especially when compared to instructionist gaming which involves the mastery of complex interfaces but rarely to reaches beyond the surface of the screen itself. By their very design, digital games provide compelling systems precisely because they are not only one of the first systems a child encounters at an early age, but they remain a regular presence in children's lives, even as they graduate into adulthood. Video games are no longer meant for children only and as players grow older, they increasingly can appreciate the nuances of and the differences between various games. The design of the interface, the intuitiveness and responsiveness of game play, the way in which challenges are scaled to grow more complex and difficult—even where there are potential "cheats" within the game engine—all can be manipulated by the player. All of these functions are present in even the most rudimentary of video games and are optimal fodder for exploring the nature of systems, particularly when a player is not simply reacting to the system but also designing it.

Numerous studies over the last twenty years have shown what students can learn when coding games (e.g., Hayes & Games, 2008) using various programming tools such as Agentsheets, Alice, Flash, Logo, Scratch among others (Burke & Kafai, 2014). In a quasi-experimental study, Kafai (1995) showed that a class of upper elementary students who learned Logo programming in the context of game design activities over a three month time period improved significantly in writing and debugging programs when compared to students who were learning Logo programming solely in the context of smaller independent projects unrelated to gaming.

Supporting this initial research, several comparative pair programming studies (Denner & Werner, 2007) showed that pairs of Latina middle school girls outperformed students working on their own in learning programming concepts when designing games with Alice. Both designing and playing a video game alongside peer proved to be a crucial way that children understood the nature and function of code. The Globaloria network (Reynolds & Caperton, 2007), in which over thousands of students design video games as part of curricular activities in their schools, also demonstrated learning of key programming concepts using Flash. Even outside of school, a two-year study in a Computer Clubhouse found that use of programming concepts significantly increased from year 1 to year 2 (Maloney et al., 2008) as children increasingly developed and remixed video games for themselves and each other.

Foremost though, coding has received by far the most attention because it can include various software design practices ranging from programming, debugging, and remixing code. Taken together, these practices capture what has been described as "computational thinking" which Wing (2006) defined as designing systems for more effective problem solving. While computational thinking is not just coding, code represents one of the key avenues to engage youth in an early understanding about how effective systems are designed and maintained, a skill set that can be applied to fields as diverse as industrial mechanics, computational biology, and marketing analytics. Understanding game design is an optimal early incubator for grasping computational thinking as would be designers not only have to create a series of novel user interfaces but also need to ensure that these interfaces scale in complexity and even adjust to the player's capacity to accomplished digitally-designed tasks. Coding in the context of constructionist gaming is not just learned for the sake of understanding and generating code, it also demands designers to be aware of perspectives other than their own and thus provides a rich context for collaboration, the next dimension.

Collaboration

The *collaborative* dimensions of constructionist gaming is often perceived in terms of the exclusive communities of "gamers" who are the self-professed experts in all things video games, much to the frequent annoyance of others and even to the isolation of themselves. Yet this die-hard group of those who "geek out" (Ito et al., 2009) around making and playing video games tend to overshadow a growing number of DIY communities that use programming as a core tool for creative media production, including robotics communities, e-textile communities, and programming communities like those that have evolved around Scratch, Arduino, and Processing languages. As large online communities have grown around more beginner-friendly tools like Alice, Scratch and Processing, they are marked by openness rather than what has traditionally been perceived as gamer exclusivity, with members regularly sharing ideas and remixing one another's work. These new tools further reshape contemporary literacy practices in DIY communities, helping youth to meet the goals of becoming fluent with technologies, and extend computational thinking into computational participation (Kafai & Burke, 2014), in which solving problems and designing systems are not solely the function of algorithmic processes but more fundamentally representative of the practices and perspectives increasingly necessary to contribute within wider social networks and understand the cultural and social nature of a networked society.

This push for more collaborative endeavors around making video games becomes readily apparent with the plethora of new gaming challenges that have grown popular just over the past three years. The aforementioned Scratch website issues regular "collaborative challenges" and "collab camps" (Kafai & Burke, 2014) annually, as does Microsoft's Kodu site with the "Kodu Cup" and Globaloria with its annual "Globey Awards" challenge. While each of these sites have their own rules and regulations for their respective competitions, all of these competitions foster the collaborative spirit by encouraging their challengers to post their ongoing projects for feedback from their peers and utilize discussion boards and forums to search out fellow team members and solicit advice on the game-making process. In the spirit of competition, collaboration (more tacitly) ensues—not unlike what we witnessed with the tremendous growth of science fairs over the second half of the 20th century in the U.S. Even the federal government appears to be tapping into the excitement of gameplay and competition, having sponsored the STEM National Video Game Challenge (http://www.stemchallenge.org) for the past three years. With the stated goal "to motivate interest in STEM learning among America's youth by tapping into students' natural passion for playing and making video games," the Challenge is issued by none other than the President himself and can be utilized as a forum that K-12 schools can adopt to more effectively integrate collaborative STEM learning through a hands-on, project driven approach.

This constructionist context of making games for others adds a new collaborative dynamic to the more traditional instructionist approach to gaming in which the power of collaboration manifested itself in players playing with each other to advance the game. For instance, Gee (2003) brings up examples of how players have to coordinate in order to plan and successfully orchestrate many of the higher up challenges that no single player could can complete on his or her own. Likewise, Luther and Bruckman (2011) illustrate how in online creative communities, such as the popular game-making site Newgrounds, when collaborations succeed, they produce

content that far supersedes what any single member could have made on his or her own. In these collaborative game-making activities, different expertise, not just technical but also team management and various creative artistic skills are needed, which leads to our next dimension.

Creative

Many of the creative practices involved in making games are rooted in the arts and can involve observing and deconstructing media, evaluating and reflecting gameplay, as well as referencing, reworking and altogether remixing other games (Hetland et al., 2007). These referencing, reworking and remixing practices include not only the creation of original works that make knowing reference to previous games, cartoons, music, and other sources of popular culture but also the modification of existing games, images or sounds, often to create entirely new interactive pieces or "machinima" such as non-interactive movies. With the advent of the so-called "web 2.0", youths' creative media production with digital media has increasingly entailed a great deal of reworking or remixing of popular media texts such as videogames and music (Kafai & Peppler, 2011).

Peppler and Kafai's (2007) case study of 15-year old Jorge well captures the potential for young game designers to not only learn coding and effective collaboration but also the creativity behind seamless imitation. Using Scratch at a Computer Clubhouse specifically geared toward low-income youth from the surrounding neighborhood, Jorge was a regular visitor to the Clubhouse over the eight-months of the ethnographic study. The second project he created was a video game entitled "Metal Slug Hell Zone X", a play off the popular "run and gun" video game series Metal Slug. Carefully coding each sprite within Scratch to respond promptly to keystrokes, Jorge fully recreated the avatar fluidity characteristic of the original game, exploring and— to a certain degree—reformulating the genre conventions of shooter games. Yet with this functionality established, Jorge did not stop. Instead, he spent numerous additional hours, drawing every character and animation from the original game using Scratch's paint feature, which in turn were based upon his own penciled sketches of the original video game. If, as Buckingham (2003) points out, "imitation is an indispensable aspect of learning" in media education, Jorge's own video game exemplifies the educational potential of such creative imitation (p. 134).

In observing creative practices as they pertain to constructionist gaming, young designers learn about and appreciate artistic principles by *making artistic choices* within a single modality (e.g., visual, audio, or kinesthetic), as well as by *connecting multimodal sign systems* across two or more modalities (e.g., visual and sound, visual and movement or gesture, and sound and movement) to convey an artistic idea (Peppler, 2013). In constructionist gaming, the creative dimension not only adds personal but also multimodal expressions into their designs. Some of these same elements of creative designs arise in instructionist gaming, such as when players have the opportunity to name and customize their avatars at the start of a game. While this may seem to be only a minor element, this ability to creatively customize a game has been one of the hallmarks of "good" instructionist games and points out that the creative dimension of serious gaming has always been inherently constructionist in nature.

Critical

The fourth and final component of constructionist gaming—criticality—may very well be the hardest to pin down since media and arts education have historically emphasized the consumption and appreciation of existing designs as their main goals rather than making anew through the critical remixing and repurposing of such designs (Peppler, 2013). Several approaches have examined game design as a way to involve youth in critically viewing media and using this understanding to create their own original work. As youth begin to take advantage of living in a digital world by capitalizing on the wealth of images, sounds, and videos accessible as "materials" to reuse in their own work, media educators grew particularly concerned about the ways in which youth are either re-inscribing or questioning existing dominant norms (Buckingham, 2003; Buckingham & Burn, 2007). These critical practices of game production include youth being able to critically reflect on and evaluate media texts, understanding references made in popular texts, and deconstructing and interpreting the meaning behind such texts. By observing the critical practices of game designers in this way, we gain an understanding of the extent to which young designers understand and question the popular texts that they incorporate in their work, apart from what they learn about software programming and the arts.

For instance, critical choices can take on the form of game designers intentionally removing all shooting features and enemies while keeping other features of a run-and-gun game genre intact (e.g., side-scrolling engine, smooth-action animation, core mechanics, etc.) to create a peaceful setting in a once violent videogame (Peppler & Kafai, 2007a). Popular DIY practices, like remixing, bring up important issues of ethics in new media literacies such as *crediting ownership* and *providing inside information*. Crediting ownership consists of referencing the intellectual origins of "text" used in media productions. And children can take this referencing quite seriously. In an after-school club, Scratch programmers ages 10-12 years were adamant that their fellow programmers credited the origins of programs that they had remixed and posted online. While Scratch programmers initially were concerned about other taking their programs, they also came to understand

the remixes as a form of recognition that represented attention they received from others (Kafai, Fields, & Burke, 2010).

In the context of constructionist gaming, the process through which youth transform from players to creators of gaming also provides a critical lens, even in informal learning spaces (Burke & Kafai, 2014; Peppler & Kafai, 2007a). The extent to which these practices represent the larger community is unknown and is at the core of our rationale for investigating vast data sources that were amassed by multiple members of the community. As youth make a series of choices, this ultimately leads to more fuller forms of literacy as they become more practiced in these decision making processes. Some approaches in instructionist gaming have taken on more critical lenses by choosing topics such "World without Oil", which engages game players in critical examination of their own gas consumption by imagining over several weeks what life would be like with restricted or even absent resources; likewise, DeVane and Squire's (2007) study on how youth of different SES play Grand Theft Auto and use this to examine home ownership and other economic situations holds this element of personal critical reflection within an instructionist gaming context. Whether within an instructionist or constructionist context though, this fourth C of criticality is a crucial element to keep in mind whenever playing or making video games as games and gaming do not simply represent an escape from everyday life but also a reflection of our own lives and personal predilections.

Connected Gaming

We see the four C's—coding, creative, collaborative, and critical practices–to be present in both constructionist and instructionist approaches to gaming, and while some of these elements are more widely documented to occur in one approach over the other, these two approaches are nonetheless complimentary and serve as the basis for connected gaming. The well-known game SimCity and the newly released Scratch 2.0 program each offer an apt example of instructionist and constructionist approaches merging together into this notion of connected gaming. From the instructionist gaming side, SimCity illustrates how playing a game can contribute to a better understanding of the constantly shifting dynamics of a simulated world (Salen, 2013). From the constructionist gaming side, new features in Scratch 2.0 environment allow for writing programs that survey information from participants at the site to better understand who is sharing online and what they are sharing (Dasgupta, 2013). These are two different approaches, but both have the same goal of "looking under the hood" for understanding what happens in the massive and interconnected community. While the tools in SimCity are programmed by experts, the tools in Scratch are programmed by players themselves. Going forward, there is no reason that SimCity couldn't offer programmable tools that would allow end-users to customize their investigations, while pre-programmed tools in Scratch can be offered for those wanting to experience an actual simulation before designing their own. In fact, the latter approach already exists. Thus in bringing instructionist and constructionist approaches together, we open up new perspectives on using computation for understanding online participation in gaming.

To realize this potential of connected gaming, however, we face at least two critical challenges that have long faced gaming culture in an instructionist context: access and participation. The first issue stems from the lack of access to learning coding skills. While children may have the devices themselves, they have little to no understanding how the devices actually work. The second issue follows the first and addresses the strong disparities in participation as to who actually produces within both gaming and coding communities. These two issues have been particularly dicey issues for girls with girls' underrepresentation in both coding and gaming communities. Yet, in an unexpected development over the last decade, programming games has been used to broaden participation in computing for girls (Kafai et al., 2008).

From Tools To Communities in Connected Gaming

So can we capitalize on these developments to broaden access and participation? Making games is obviously not a simple enterprise but requires much, including dedicated chips, significant technical knowledge, as well experience in storytelling, art, and design. Can novices become such game designers? One of the key challenges is to provide them with tools that lower the barriers, or the floors (Burke & Kafai, 2014), to make the once-laborious process of computer programming. But by the same token, with the floor laid out, the next challenge becomes to what extent these various game-making tools have the capacity to *retain* their users. While "low floor" accessibility is the first step to ensure a steady number of novice users are accessing and using a game tool, designers also have to ensure that their game engine is robust enough to ensure more experienced users do not tire of the software and can find new ways to become more proficient at making video games..

But most importantly, tools also need to consider the participation issue and with it, shift their attention to the larger gaming community. Here "wide walls" signify the capacity of a tool to allow for a variety of creations—in this case, a wide variety of games. Effective game-making tools must allow their users to create a variety of game genres, be it platform games, first-person shooter games, RPGs (role-playing games), strategy games, and trivia games, to name a few. Likewise, "large windows" provide opportunities to connect with others to join gaming communities that revolve around same interests. Many communities now are connected to

game design tools, including Kodu, Scratch, GameStar Mechanic, Spolder, and Game Salad. Some of these are specific to video games while others are more open-ended and allow for multiple designs besides games. These communities of game designers is a key component, from the early classrooms where kids designed their individual games to the massive online communities where games are some of the most popular designs shared. Gaining access to a wide and appreciative community means that players have the opportunity to leverage that community as an extension of the tool itself, with meaningful feedback serving to help fledgling designers gain a foothold into what works in game design, while more experienced designers can grow in proficiency and create increasingly intricate games.

From Old to New Clubhouses in Connected Gaming

If tools can provide access to new communities and communities can function as effective extensions of such tools, then we also need to address who is participating and who can participate in these communities. Gaming (Jenkins & Castell, 1998), but also coding communities (Margolis & Fisher, 2002), have a long history for not engaging girls and the reasons are multiple: on one hand, there is the lack of interest, lack of experience, and lack of skill from females, while on the other hand there is the persistent stereotyping of women in these same three areas, which is then compounded by a lack of female player roles and the prevalence of violence in games. This larger issue of gender differences is not germane to gaming alone—it is one that has plagued programming and STEM in the learning sciences at large. And yet, despite these persistent issues, constructionist gaming approaches have been seen as a possible remedy for addressing the gender divide so present in the gaming culture at large. An early study of game making revealed no significant gender differences in learning programming and disbanded with conventional wisdom at the time believed to be true: girls could be interested in programming and be interested in gaming, if they were just given the opportunity to make their own (Kafai, 1995). The success of girls in constructionist gaming became the launch pad for a whole series of tool developments (such as Storytelling Alice) and research initiatives to use game design to broaden girls' participation in computing.

While there was much success with game making to bring girls into the so-called clubhouses of computing and gaming, it also revealed a problematic aspect: why did girls have to design games in order to become gamers and more tech-savvy? This issue received little attention, even from the feminist side who mightily and justifiably lamented about the reification of stereotypes in girls making games (Jenson & deCastell 2007). The challenge we are faced with is to no longer simply question how to open the doors of existing technology and gaming clubhouses but how to build new clubhouses that envision different applications and activities in computing and gaming. The most prominent example here is the work by Leah Buechley who re-designed the Arduino board into the LilyPad Arduino for making electronic textiles. She found, indeed, new communities or clubhouses of coding could be created with such redesigns that are functionally equivalent in their technical complexity but application-wise result in the construction of different artifacts (Hill and Buechley, 2011). A possible equivalent in building new clubhouses for gaming could be to focus on the relationship between stories and games and conceiving of the game making process as a matter of crafting pathways rather than simply responding to stimuli (Westecott, 2012).

We of course have only touched upon the surface in imagining what connected gaming could look like and how it can begin to address these issues of access and participation. When the field of serious gaming started, attention nearly inordinately focused on proving the effectiveness of instructionist gaming (Clark, 2007) and "researching learning in popular gaming cultures, designing learning environments based on those principles, and reconceptualizing educational practice for an interactive age" (p. 51, Squire, 2007). Constructionist gaming really was not part of either discussion in building the field of serious gaming. But if we want to realize the larger potential of serious gaming, we need embrace a larger agenda that recognizes that opening access and participation in serious games is not solely a matter of making better games for the end user but allowing these end users themselves to make the games they would like to see and play. Ultimately, connected gaming's goal is to promote environments good for learning, and it is here where constructionist approaches join instructionist efforts. This is the case for "connected gaming", an approach that doesn't draw boundaries between players and designers as participants of digital media culture but rather sees them as complimentary to each other as already Papert envisioned: "if one does belong to a culture in which video games are important, transforming oneself from a consumer to a producer of games may well be an even more powerful way for some children to find importance in what they are doing" (p. iii, 1995).

References

Buechley, L. & Hill, M. (2012). LilyPad in the wild. How hardware's long tail is supporting new engineering and design communities. *DIS '10 Proceedings of the 8th ACM Conference on Designing Interactive Systems* (p. 199-207) New York, NY: ACM

Buckingham & Burn (2007). Game literacy in theory and practice. *Journal of Educational Multimedia and Hypermedia*, *16*(3), 323-349.

Burke, Q. & Kafai, Y.B. (2014). A decade of programming games for learning: From tools to communities. In H. Agius & M.C. Angelides (Eds.) *The Handbook of Digital Games: Institute of Electrical and Electronics Engineers (IEEE)*. New York: John Wiley & Sons.

Burke, Q. & Kafai, Y.B. (accepted). DIY zones for Scratch designs in class and club. *The International Journal of Learning and Media. 3(4).*

Clark, R. E. (2007). Learning from serious games? Arguments, evidence, and research suggestions. *Educational Technology, 47*(3), 56-59.

Cuban, L. (1986). *Teachers and machines: The classroom use of computers since 1920*. New York, NY: Teachers College Press.

Dasgupta, S. (2013). From surveys to collaborative art. *Proceedings of the 12th International Conference on Interaction Design and Children* (pp. 28-35). New York, NY: ACM.

Denner, J., Werner, L., Bean, S., & Campe, S. (2005). The girls creating games program: Strategies for engaging middle school girls in information technology. *Frontiers: A Journal of Women's Studies*, *26*(1), 90-98.

Denner, J. & Werner, L. (2007). Computer programming in middle school: How pairs respond to challenges. *Journal of Education Computing Research*, *37*(2), 131-150.

DeVane, B. & Squire, K. (2008). The meaning of race and violence in Grand Theft Auto San Andreas. 10.1177/1555412008317308.

Duncan, S. (2011). Minecraft, beyond construction and survival. *Well Played, 1*(1), 1-22. New York, NY: ACM.

El Nasr, M. & Smith, S. (2006). Learning through game modding. *Computers in Entertainment, 4*(1) 10.1145/1111293.1111301.

Gee, J.P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gentile, D. A., & Gentile, J. R. (2008). Violent video games as exemplary teachers: a conceptual analysis. *Journal of Youth and Adolescence, 37(*2), 128-141.

Girard, C., Ecalle, J., & Magant, A. (2012). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *(29)*3. 1-13.

Grover, S. & Pea, R. (2013). Computational Thinking in K–12: A Review of the State of the Field. *Educational Researcher, 42* (1), 38-43.

Hayes, E.R. & Games, I.A. (2008). Making computer games and design thinking: A review of current software and strategies. *Games and Culture*, *3*(4), 309-322.

Casell, J. & Jenkins, H. (1998). *From Barbie to Mortal Kombat*. Cambridge, MA: MIT Press.

Jenkins, H. Clinton, K., Purushotma, R., Robison, A., & Weigel, M. (2006). Confronting the challenges of participatory culture: media education for the 21st century. White Paper. Chicago, IL: The John D. & Catherine T. MacArthur Foundation.

Jenson, J. & deCastell, S. (2007). Girls playing games: Rethinking stereotypes. *Futureplay*, 9-16.

Kafai, Y. B. (1995). *Minds in play: Computer game design as a context for children's learning*. Mahwah, NJ: Lawrence Erlbaum.

Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. *Games & Culture*, *1*(1), 36-40.

Kafai, Y.B. & Burke, Q. (2014). *Connected code: Children, programming and the power to make and connect.* Cambridge, MA: MIT Press.

Kafai, Y. B. & Fields, D. A. (2013). *Connected play: Tweens in a virtual world*. Cambridge, MA: MIT Press.

Kafai, Y.B., Fields, D.A., & Burke, W.Q. (2010). Entering the clubhouse: Case studies of young programmers joining the online Scratch communities. *Journal of Organizational and End User Computing*, *22*(2), 21-35.

Kafai, Y. B. & Peppler, K. A. (2011). Youth, technology, and DIY: Developing participatory competencies in creative media production. *Review of Research in Education, 35*(1), 89-119.

Kelleher, C. & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys, 37*(2), 88-137.

Klopfer, E. & Haas, J. (2012) *The more we know: NBC News, educational innovation, and learning from failure.* MIT Press, Cambridge, MA.

Knobel, M. & Lankshear, C. (Eds.) (2010). *DIY media: Creating, sharing, and learning with new technologies*. New York: Peter Lang.

McLaurin, M. (2009). Kodu: End-user programming and design for games. In *Proceedings of FDG '09.* New York, NY: ACM. DOI: 10.1145/1536513.1536516

Mote, C., Kafai, Y.B., & Burke, Q. (2013). Can online competitions be collaborative? Utilizing virtual events to promote communication and learning in schools. *Learning & Leading with Technology, 41*(4), 16-21.

National Research Council (2011). *Learning Science through Simulations and Games*. Washington DC: The National Academies Press.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books. New York.
Pelletier, C. (2008). Producing difference in studying and making computer games: How students construct games as gendered in order to construct themselves as gendered. In Y. B. Kafai, C. Heeter, J. Denner, & J. Sun (Eds.), *Beyond Barbie and Mortal Kombat* (pp. 145-161). Cambridge MA: MIT Press.
Peppler K. & Kafai, Y.B. (2007a). From SuperGoo to Scratch: Exploring creative digital media production in informal learning. *Learning, Media, and Technology*, *32*(2), 149-166.
Peppler, K. Kafai, Y. B. (2007b). What video game-making can teach us about learning and literacy: Alternative pathways into participatory culture. In Akira Baba (Ed.), *Situated Play: Proceedings of the Third International Conference of the Digital Games Research Association (DiGRA)* (pp. 369-376). Tokyo, Japan: The University of Tokyo.
Repenning, A. & Ioannidou, A. (2008) Broadening participation through scalable game design. *SIGCSE '08 Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education* (pp. 305-309). New York, NY: ACM.
Resnick, M., Maloney, J., Monroy Hernandez, A. Rusk, N., Eastmond, E., Brennan, K. Millner, A., Rosenbaum, E., Silver, J., Silverman, B. & Kafai, Y.B. (2009). Scratch: Programming for all. *Communication of the ACM, 52*(11), 60-67.
Reynolds, R. & Harel Caperton, I. (2011). Contrasts in student engagement, meaning-making, dislikes, and challenges in a discovery-based program of game design learning. *Educational Technology Research and Development*, 59, 267-289.
Salen, K. (2007). Gaming literacies: A game design study in action. *Journal of Educational Multimedia and Hypermedia, 16*(3), 301-322.
Salen, K. (2013, March). Agile development meets evidence-centered design: Glasslab and the design of game-based assessments. Presentation at the Digital Media & Learning Conference, March 14-16. Chicago, IL.
Shaffer, D. (2007). *How computer games help children learn*. New York, NY: Palgrave MacMillan.
Steinkuehler, C. & Duncan, S (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education Technology*, 17(6), 530-543.
Squire, K. (2007). Games, learning, and society: Building a field. *Educational Technology*, *47*(5), 51-55.
Squire, K. (2011). *Video games and learning: Teaching and participatory culture in the digital age*. New York: Teachers College Press.
Torres, R.J. (2009). Using Gamestar Mechanic within a nodal learning ecology to learn systems thinking: A worked example. *International Journal of Learning and Media*, *1*(2), 1-8.
Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2013). Computer gaming and interactive simulations for learning: a meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.
Westecott, E. (2012). Independent game development as craft. *Loading, 7*(1), 78-91.
Wouters, P., van Nimwegen, C., von Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 1-17.
Young, M. F., Slota, S., Cutter, A. R., Jalette, G., Mullin, G., Lai, B., Simenoni, Z., Tran, M., & Yukhymenko, M. (2012). Our princess is another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82*(1), 61-89.

Acknowledgments

The writing of this paper was supported by a collaborative grant from the National Science Foundation (NSF-CDI-1027736) to Yasmin Kafai, together with Mitchel Resnick and Yochai Benkler. The views expressed are those of the authors and do not necessarily represent the views of the Foundation, the University of Pennsylvania or the College of Charleston.

Learning with Multiple Visualizations in the Science Museum

Joyce Wang, University of Pennsylvania, joyce.s.wang@gmail.com
Susan Yoon, University of Pennsylvania, yoonsa@gse.upenn.edu

Abstract: Science museums are intentionally designed spaces that foster visitors' understanding of scientific knowledge. Increasingly, museums are adopting digital media and technologies in the exhibits both to modernize the experience and to increase visitors' interest, engagement, and learning. This study examines how three dynamic visualizations (digital augmentation, computer simulation, and animation) support visitors' knowledge of a commonly misunderstood scientific concept, Bernoulli's Principle. Data from interviews, surveys, and tests reveal that visitors' knowledge significantly increased after engagement with multiple visualizations. Both children and adults attributed their understanding to the affordance of multiple visualizations to accommodate a range of learning styles and to offer a diverse range of types and depth of knowledge. Based on these findings, we suggest that designing for multiple visualizations in museum exhibits is a positive approach to increasing visitors' understanding of scientific knowledge.

Introduction

Informal environments such as museums play a prominent role in our nation's science education landscape. Intentionally designed to support learning about the physical and natural world around us, research has found that science museums foster engagement and interest in science, cultivate the understanding of scientific knowledge, support the growth of scientific reasoning skills, encourage reflection on science, promote engagement in scientific practices, and advance the development of science learner identities in visitors (NRC, 2009). In particular, understanding science knowledge, such as concepts, facts, models, and explanations, is an important motivation for many museum visitors including teachers bringing school groups (Kisiel, 2005) and parents bringing their families (Falk & Storksdieck, 2010). Similarly for museums as educational institutions, being able to accurately represent and successfully communicate important scientific concepts to enhance the general public's science understanding is an important institutional goal. While several studies document that museum visits enhance visitors' science knowledge understanding, assessments that measure this knowledge often demonstrate little or no positive change in science knowledge outcomes for learners (NRC, 2009).

The question of how best to support learning in science museums is often related to exhibit design. Because museums lack the direct facilitation, accountability, and rigid structure that characterize formal learning environments, supporting free-choice learning that is based on visitors' interests and motivations requires intentional design of museum objects, labels, and spaces. Increasingly, museums are adopting new digital technologies to support science learning. A growing body of literature argues that these technologies can contribute positively to visitor experiences. For example, Sandifer (2003) found that visitors tend to use technology-based exhibits more frequently and for longer periods of time than traditional exhibits, Laursen (2013) presented illustrations of children engaged in various levels of co-participation around a computer-based device, and Eberbach and Crowley (2005) found that virtual representations of objects support different kinds of learning. However, concerns have also arisen about the negative effects of technology on visitors, such as the tendency for them to interact less with other exhibits or objects (Ucko & Ellenbogen, 2008). "Ultimately, the goal of introducing new media technologies…is not only to modernize the experience and space, but to significantly improve the quality of the visitor experience, including enhancing learning outcomes" (NRC, 2009). Fundamentally, more research needs to analyze the effects of digital technologies on museum learning.

It is within this field of research that we now position our study. In this paper, we investigate how various digital technologies impact visitors' knowledge of Bernoulli's Principle, a concept often illustrated in science museums. Specifically, we examine how dynamic visualizations can support science learning. Drawing upon Ainsworth's (1999, 2006) work on multiple representations, we study how the combination of three visualizations (digital augmentation, simulation, and animation) together can afford learning of scientific concepts, facts, and principles. The research question investigated in this article is "How do multiple visualizations enhance learning in a science museum?"

Theoretical Considerations

Exhibit design is a critical feature of visitor engagement and learning. From the makeup of the individual interpretive labels to the arrangement of the exhibit elements on the museum floor, careful attention is directed towards how exhibit features might attract visitor attention and facilitate learning. In our earlier work, we investigated how scaffolding the physical design of the exhibit device might impact school children's learning. We examined how the addition of knowledge-building scaffolds and digital augmentation might enhance

learning, and we discovered that knowledge-building scaffolds support cognitive learning and digital augmentation supports conceptual understanding (Wang & Yoon, 2013; Yoon & Wang, 2014; Yoon et al., 2012a, b, 2013b). Furthermore, we found that digital augmentation supports conceptual learning because it encompasses many of the same learning affordances as dynamic visualizations (Yoon & Wang, 2014). Building upon these earlier findings, this study considers how adding more dynamic visualizations might support even deeper learning during museum visits. Briefly, the device under investigation, *Bernoulli Ball,* depicts a lightweight plastic ball that is able to float in the air due to the interactions between the speed and pressure of two types of air – the normal air in the room and the air that is being blown out of a blower attached to the device. In this section, we'll first briefly discuss some of the general learning affordances of dynamic visualizations to set the framework for understanding why they can be beneficial to learners. We'll then address the three specific visualizations that were employed in this study and conclude by presenting Ainsworth's work on multiple representations to help ground our rationale for using multiple visualizations in museum learning.

Dynamic Visualizations

Dynamic visualizations, or external representations that are able to depict changes in space over time and a continuous flow of motion, have become a popular means of providing instruction in all types of learning environments (e.g., Lowe & Schnotz, 2008). There are four main affordances of dynamic visualizations that make it so attractive for learning. These include attracting learners' attention and motivation (Scheiter et al., 2009), enabling the visualization of invisible entities or processes (Hegarty, 2004), allowing objects to be viewed from different angles or viewpoints (Hegarty, 2004), and increasing interactivity and control which can facilitate comprehension (James et al., 2002 as cited in Plass et al., 2009). Collectively, these affordances significantly aid thinking.

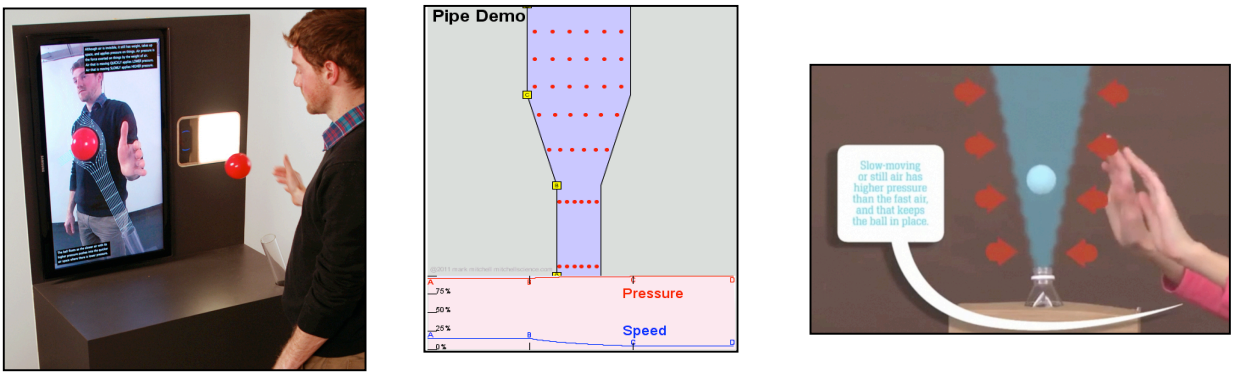


Figure 1. Three visualizations: digitally augmented device, simulation, animation.

Three dynamic visualizations were investigated in this study. The first was digital augmentation, which we previously defined as computer-generated images superimposed upon the physical environment (Yoon et al., 2012b). Although its use is relatively new within museum spaces, some studies have revealed that they can elevate visitors' interest and engagement (Hall & Bannon, 2006), support collaborative interactions (Asai et al., 2010), and garner conceptual understanding (Yoon et al., 2013b). For this study, the digital augmentation (Fig. 1) depicts the movements and pressures of the two types of airs using different colored arrows. These arrows would readjust their position depending on where the ball was. The second visualization employed was a computer simulation. Computer simulations are programs that model a simplified system or process of a real-world phenomenon and their ability to allow users to manipulate variables and observe the resulting changes makes them effective visualizations (de Jong, 2011). In science museums, they have been presented as fun games that invite visitors to explore certain aspects of a particular scientific issue (e.g., Cheng et al., 2011). The simulation in this study, depicted in Fig. 1, shows a pipe with molecules flowing through and a line graph below illustrating the speed and pressure changes of these molecules. Individuals were able to manipulate the width of the pipe and examine its effects on the pressure and speed of the flowing molecules. The last visualization employed was an animation. Animations are "pictorial display that changes its structure or other properties over time and which triggers the perception of a continuous change" (Schnotz & Lowe, 2008, p. 34). The ability to visually depict changing information, some of which are not observable is one of the greatest benefits of using animation (Rogers, 2008). In museums, animations have been used as virtual staff members that guide visitors in particular ways (Lane et al., 2011). They've also been used to convey important scientific information; consequently data has revealed that visitors respect these animations as trusted sources of information (Matuk & Uttal, 2008). The animation (Fig. 1) in this particular study demonstrates how the phenomena can be replicated at home with a blow dryer and a Ping-Pong ball. A child narrates the set-up and briefly describes the science behind how the ball floats.

Multiple Representations

As discussed above, whether through increasing learners' interests to continue their exploration or by making information more visible and accessible, dynamic visualizations bear much potential to support learners' comprehension of complex scientific phenomena. In this study, we hypothesize that by bringing together these individual visualizations, we'll be able to capitalize on their collective affordances to garner even deeper learning than if using them separately. We draw upon Ainsworth's work on multiple external representations (MERs) to provide some grounding on which to base this hypothesis.

MERs have been widely used in science teaching to help learners understand complex scientific concepts (Ainsworth, 1999). They typically refer to modern, multi-representational, computer-based learning environments that package together several dynamic representations such as audio, video, animations, dynamically changing graphs, diagrams, and tables and other interactive dynamic visuals (van der Meij & de Jong, 2006). Although a common justification for employing multiple representations is that they are more likely to capture a learner's interest and motivation (Ainsworth, 1999), they also provide cognitive benefits that aid in learning. Broadly, there are three main functions of multiple external representations: to complement each other, to constrain the interpretation between the visualizations, and to construct deeper understanding (Ainsworth, 1999). First, MERs can complement each other in the content of the representation or in the cognitive processes needed to interpret the representation. Concerning content, regardless of whether they express completely different information or they provide some similar redundant information, by distributing the information across several representations instead of containing all of it in a single representation, the complexity of the representation decreases. This in turn decreases the amount of cognitive load learners need to interpret the representation (Ainsworth, 1999). MERs can also support complementary processes; even if representations contain equivalent information, they can still support different inferences because of differences in their computational properties (Ainsworth, 1999). This requires that different cognitive processes are needed to interpret the information. Second, MERs can also be used to constrain the interpretation between the representations (Ainsworth, 1999). A familiar representation can be used to support the interpretation of a less familiar or more abstract representation. In this role, the familiar representation is meant to support learners' reasoning about the less familiar one. The last function of MERs is to support the construction of deeper understanding by promoting abstraction, encouraging generalizations, and in teaching the relations between representations (Ainsworth, 1999).

In this section, we've outlined how previous studies on dynamic visualizations and MERs inform our hypothesis of bringing together several dynamic visualizations to enhance learning. Because our research analyzes the use of three distinctive visualizations that operate through three different platforms, from here on out, we will use the term "multiple visualizations" (as opposed to MERs) to refer to the collective of the three distinct visualizations.

Methods

Context and Participants

This study evolved from a formerly funded large-scale National Science Foundation informal science education project in which the goal was to design, integrate, and increase the use of educational technologies, particularly digital augmentation, and to study their impact within the science museum learning experience (Yoon et al., 2012a, b, 2013a). This mixed methods quasi-experimental study extends that focus by considering how two additional technologies, a computer simulation and an animation, can supplement the previously augmented exhibit device. The participants were family groups that consisted of at least one child, between the ages of 11 and 14, and one adult. In total, 30 families with 37 children and 38 adults participated. Over 75% of the families were recruited directly off the museum floor on the day of the study. The remainder was recruited through the museum's monthly community outreach event and through emails to professional and personal contacts.

This study follows a within-subjects, or repeated measures, design in which all of the participants engaged in both experimental conditions with the device. The study took approximately 45 minutes to complete and was held in a separate room off of the museum floor. Before families engaged with the exhibit device, parents were asked to fill out a pre-survey and children were asked to complete a pre-knowledge test and answer some pre-intervention interview questions. After this pre-intervention data was collected, families were exposed to the first condition that featured the device with just the digital augmentation visualization. Families were asked to play with the device as if they had found it on the museum floor. Once families signaled that they were finished playing and after the mid-intervention interview data was captured from children, they moved onto the second condition. The second condition featured a computer simulation and an animation in addition to the augmented device. Both the simulation and animation were presented on two netbooks at a table adjacent to the device. Again, families were asked to play with all of the tools as if they had seen it on the museum floor and to signal me when they were done. Their participation ended once parents completed their post-survey, children

finished their post-knowledge test and post-intervention interview, and both children and parents answered some post-study interview questions.

Data Sources and Analysis

1. Pre- and post-surveys were administered to adults at the very beginning and end of the study. There were two parts to the surveys but for the purposes of this study, only the second part, which assesses parents' *knowledge of the phenomenon*, was analyzed. The question asks, "How do you think the ball is able to stay floating in the air without being blown away?" In the post-survey, we asked parents to review and revise their answer. A categorization manual was constructed to evaluate this question and interrater reliability was obtained on 20% of the responses by two independent researchers ($\alpha = 0.91$). The written responses were coded from "Little to No Understanding" (Level 1) to "Complete Understanding" (Level 6). Whereas in a level 1 response, the individual attributes the floating phenomenon to just the air from the blower, a level 6 response recognizes that the ball floats because the high pressure from the slow-moving room air keeps the ball in the low-pressure, fast-moving air stream. Both types of air are explicitly addressed in the highest level of understanding. Paired samples *t*-tests were conducted to determine whether parents' understanding of the phenomenon had changed after the study.
2. Children's *knowledge of the phenomenon* was also measured, but was administered via interview as opposed to a written survey. Pre-, mid-, and post-intervention interviews were conducted with children to more finely explore how their knowledge changed with each exposure to the visualization tools. In all three interviews, children were asked the same question as on parents' surveys, "How do you think the ball is able to float without being blown away?" The mid-intervention interview was administered to more finely identify differences in the learning impact between just the digital augmentation and then the multiple visualizations. The same categorization manual was used to code these verbal responses. Interrater reliability was obtained on 20% of the responses by the same independent researchers ($\alpha = 0.96$) and a one-way repeated measure ANOVA was conducted to examine how their knowledge changed between each condition.
3. Pre- and post-knowledge tests were administered to children at the very beginning and end of the study to assess children's *conceptual understanding*. Whereas the interviews measured their understanding of the exhibited phenomenon, these tests were meant to assess understanding and application of the principle to different contexts. The test consisted of 4 multiple-choice questions (2 low level recall questions and 2 application questions) that were informed by textbooks and vetted by content experts. The responses were coded as correct or incorrect and a paired samples *t*-test was administered to examine differences between overall pre- and post-knowledge scores.
4. To understand how the multiple visualizations impacted visitors personally, in the post-study interview, children and parents were asked to reflect on their interaction with all three visualizations. They were asked, "Do you think you learned more by playing with all three, or a combination of these, tools? If so, why or how did playing with these tools enhance your learning?" These responses were qualitatively mined for common themes.

Results

Adults' Knowledge of Phenomenon: Pre- and Post-Surveys

A paired samples *t*-test was performed on adults' written responses to the question "How do you think the ball was able to float in the air without being blown away?" Results indicated significant increases in their knowledge between the pre-intervention ($M = 1.87$, $SD = 0.99$) and post-intervention ($M = 3.21$, $SD = 1.61$); $t(37) = -4.97$, $p = .00$. This suggests that their engagement with the multiple visualizations in the exhibit enhanced their understanding of how the phenomenon occurred.

Children's Knowledge of Phenomenon: Children's Pre-, Mid-, and Post-Intervention Interviews

A one-way repeated measures ANOVA was conducted to compare the effects of the various visualization tools on children's understanding before any exposure to the exhibit, after the augmentation-only condition, and after the multiple visualizations condition. The results suggest that there was a significant difference in children's knowledge ($F(2,72)) = 25.399$, $p < .01$, partial $\eta^2 = 0.414$) between pre-intervention ($M = 2.16$, $SD = 1.04$), mid-intervention ($M = 2.76$, $SD = 1.52$), and post-intervention ($M = 4.27$, $SD = 1.87$). This implies that children's knowledge of the phenomenon deepened with each successive engagement with the visualizations. Furthermore, the difference between mean scores suggests that children learned most when they played with all of the visualizations.

Children's Conceptual Knowledge: Pre- and Post-Knowledge Tests

The pre- and post-knowledge tests contained two types of questions – questions about subject matter content that required simple recall of information ("What is the relationship between the speed and pressure of moving air?") and transfer questions that required application of the concept to a new situation ("What do you think will happen to the 2 hanging ping-pong balls when the boy blows air through the straw between the balls?"). A paired samples *t*-test was performed on children's test scores before and after their engagement in the mini-exhibit. Results indicated significant increases in knowledge between the pre-intervention (M = 0.84; SD = 0.83) and post-intervention (M = 1.84; SD = 1.14), $t(36) = -4.71$, $p = .00$. This similarly indicates that playing with multiple visualizations positively impacted children's conceptual knowledge of the science principles that underlie the floating ball.

Children and Adult Post-Intervention Interview about Effects of Multiple Visualization Tools

Interview responses to "Do you think you learned more by playing with all 3, or a combination, of these tools?" indicated that 100% of the families had at least one member who thought a combination was helpful while 93% of the families had every member agree that it was helpful. For example, one child remarked, "Can't just have one of it. It's not going to be enough." Parents made comments such as, "All 3 together makes it more better and easier to understand" and "I didn't understand it fully until I did all 3."

When asked to explain why or how they thought multiple visualizations supported deeper learning, two major themes emerged. First, adults and children identified that having more visualizations accommodated a greater range of learning styles. For example, one child commented "I think they're all teaching the same thing but they're all slightly different so it's kind of just what you prefer. If you prefer just listening to something, then you can do the video or if you prefer hands on, then you could do that one or kind of both *[sic]*, you could do that." Similarly parents explained, "I like when there's multiple ways…different people have different learning styles…I think maybe *[sic]* would provide a learning experience for different learning styles" and "I think because we all learn differently, the hands-on, the visual, the video....that just helps reinforce the knowledge... the person who learns most by doing or the person who learns by reading...so I think it's great, all 3, and technology today, especially for young people. I think it hit everybody". Another theme that emerged as to why engaging with multiple visualizations enhanced learning was because they offered visitors a range of types and depth of information that could be learned. Even though the visualizations all focused on the same concept, they addressed the concept from different perspectives. As one parent explains, "It's 3 different things but the same topic, same main idea on all of them...so you get to learn more about that piece of information." This affordance enabled all visitors, with varying degrees of knowledge on the topic, to gain understanding. Another parent explained, "You can go to the simulator and kind of get a more in depth look at air pressures um, maybe for kids that are ready to take it to another level of understanding...And then again, the video gave an opportunity to give kids an idea to try it on their own at home which would add another level of understanding of the experiment." This recognition that multiple visualizations could support varying depths of learning was understood to be a function of the nuanced information contained in each visualization. For example, one child explained, "I think it [enhanced learning] because it added stuff that maybe the other tools didn't have." He explains further by describing the different knowledge revealed by each tool: "Like the 3rd one. The video showed you that you could make it yourself. The simulation showed that you could change the shape and change the pressure, the speed of air by doing so. And then the augmentation showed you where the air was going. I like the 3 put together." Her brother added, "Yeah because some of the things like added little details that some of the other stuff didn't have." Here, the children have recognized that each tool presents slightly different information. The interpretation is that that when these disparate pieces of information are added together, they extend learning.

Discussion and Implications

One of the many goals of science museums is to increase visitors' understanding and knowledge of scientific concepts. Through careful design of exhibit spaces, tools, and objects, exhibit developers create experiences that not only provide entertainment but that also facilitate science learning in visitors. Consequently, this project sought to investigate how incorporating visualization tools into exhibits can support children's understanding of a commonly misunderstood scientific concept, Bernoulli's Principle. Building upon the positive learning gains from our earlier research with just one visualization tool (digital augmentation) (Yoon et al., 2012b), we hypothesized in this study that the addition of more visualizations might elicit even greater learning gains.

Overall, we found that engagement with multiple visualizations in a science museum supports science knowledge understanding. Differences between pre-, mid-, and post-evaluations of how adults and children understood how the ball was able to stay floating revealed that both groups of visitors grew in their understanding of the role of invisible features involved in the phenomenon. Initially, parents' and children's conceptions reflected an "Emergent Understanding" (Level 2) of how the system worked. Only obvious features (e.g., characteristics of the ball and the air being blown from the tube) were identified as contributing factors to

the system. These pre-intervention understandings are consistent with extant literature that documents robust difficulties in reasoning about air pressure. For example, Engel Clough & Driver (1985) demonstrated that children (between the ages of 12 and 16 years old) incorrectly associate pressure or force with movement. Similarly, Basca and Grotzer (2001) found that children often do not think that pressure exists when they can't easily see an effect or movement and Sere (1982) found that children (between 11 and 13 years old) could not imagine pressure without movement associated with it. Because air pressure is a non-obvious variable that cannot be sensed directly, we expected that children would have difficulties perceiving its role. That adults also held onto these naïve understandings was unexpected, though not completely unsurprising given the robust research on the persistence of preconceptions among older children and adults even after encountering experiences and models that contradict naïve understandings (NRC, 2000). Regarding children's understanding, we were encouraged to find positive results on the mid-intervention interviews after they had engaged with one visualization (digital augmentation), which confirms our previous research that certainly, the presence of the augmentation significantly enhances children's understanding (Yoon et al., 2012b). As digital augmentation embodies many of the advantageous qualities of dynamic visualizations in general (Yoon & Wang, 2014), its added benefits to learning is unsurprising. More importantly, we found pronounced knowledge gains in both groups of visitors after their engagement with the multiple visualizations. Whereas parents (in their post-surveys) attained a "Partial Understanding" (Level 3) of the phenomenon, children (in their post-intervention interviews) progressed to a "Basic Understanding" (Level 4). Conceptually, this indicates that both are shifting their understanding from purely obvious features (e.g., characteristics of the ball and the air being blown from the tube) to more imperceptible elements such as the interactions of other forces unrelated to the tube air (e.g., gravity or the normal air in the room pushing up/down) (Level 3) and to the recognition that these forces exert different amounts of pressure which affords the ball to float (Level 4). This conceptual growth was also evident, albeit to a lesser extent, in children's post-knowledge test. Other studies lend support to our positive findings. For example, van der Meij and de Jong (2006) found that students who were exposed to separate, non-linked visualizations (in a Physics unit titled "Moment") demonstrated significantly increased post-scores on questions about content knowledge and Ainsworth (2006) suggests that because multiple representations have the potential to support deeper understanding when learners integrate all of the information together, the insight achieved increases the likelihood of being transferred to new situations.

Based on a) our post-intervention interviews with adults and children and b) the differences between children's mid- and post-interview about the phenomenon, we argue that the growth in understanding is most likely due to their engagement with multiple visualizations in the second condition as opposed to the single visualization in the first condition. Our claim is well grounded in studies that have highlighted various advantages of learning with multiple representations. Consistent with Ainsworth's (1999) work, we found that our multiple visualizations support complementary cognitive processes. This advantage allows for learners who exhibit different preferences to exploit different visualizations according to their experiences, expertise, or familiarity (Ainsworth, 1999). Our excerpts revealed that both children and parents perceived the value of aligning learning preferences with tools that support these preferences. Several visitors highlighted the affordance of these visualizations to teach individuals who preferred learning through interactive experiences versus those who preferred learning through more direct means. Not only did they recognize that their family members learn differently, but that in spite of these differences, because of the various properties particular to each of the visualizations, every member was still enabled to learn. This affordance of multiple visualizations to accommodate particular learning styles is particularly beneficial, though there are some researchers that argue that successful learning with multiple visualizations is less about alignment to learning style preferences and more about expertise with particular subjects or representations (Ainsworth, 2006). While we would not disagree that learning with multiple visualizations corresponds to both the learner's command of the subject matter and how well s/he can interpret a particular type of visualization, we also want to articulate that these studies occurred in formal learning environments, vast contrasts to informal, museum spaces. Without direct facilitation from a teacher or prescribed assignments and activities to complete, museum learning is heavily contingent upon visitors' own personal choices, motivations, and preferences in deciding which exhibits to interact with (NRC, 2009). Several studies, including Borun and Dristas' (1997) piece on exhibit characteristics that facilitate multimodal learning, have explicitly addressed this relationship between exhibit design and visitor choice. Thus designing exhibits to optimize visitor learning must consider visitors' learning style preferences.

We also found evidence of multiple visualizations supporting complementary information. When visualizations contain partially redundant information, it enables users to exploit differences in the information, which therein supports the construction of new interpretations of the original concept (Ainsworth, 1999). In comparing the information that each visualization presented, most visitors identified that the animation taught them how to construct a similar device at home, the simulation illustrated the idea that speed and pressure is not static but rather can be manipulated, and the augmentation revealed the precise movements of the various airs involved. These pieces of information, though seemingly disparate, support the development of cognitive connections about Bernoulli's Principle when packaged together in the context of the exhibit. Visitors

recognized that despite the fact that all three visualizations addressed the same concept albeit from a different angle, the variation in details afforded by each visualization complemented each other in such a way that it deepened their learning. Packaging "elements" within an exhibit to enhance conceptual coherence is not a new practice. In fact, museums frequently cluster groups of conceptually related exhibits to communicate a main concept (Falk, 1997). However, whether visitors actually discern the underlying messages and themes across connecting exhibits has received mixed reviews (Allen, 2004; Falk, 1997). Particularly with regards to illustrating abstract concepts, visitors often have more difficulty perceiving connections and themes between related exhibits (Allen, 2004). To ameliorate this complexity, researchers and exhibit designers have called for more critical consideration to be paid to all levels of exhibit design, from small-scale user design functions to larger-scale decisions about the layout and orientation of the physical environment (Allen, 2004). In addition to incorporating explicit labels that describe the main message of an exhibit (Falk, 1997), we suggest that employing multiple visualizations that contain partially redundant information in various formats may also be an effective way to support conceptual understanding.

Much research has been devoted to understanding how museums can design for a range of learning experiences from general exhibit designs that attend to a variety of learning styles and levels of knowledge (Broun & Dristas, 1997) to addressing specific design features that influence science knowledge understanding such as labeling (Falk, 1997) and interactivity (Allen, 2004). With the advent of digital technologies in museum spaces, studies of how they impact science learning are still emerging. While some have raised the concern that technology can reduce visitors' interactions with exhibit objects or other visitors (Ucko & Ellenbogen, 2008), our study not only supports the few that have found benefits of technology on visitor learning (e.g., Asai et al., 2010; Cheng et al., 2011; Sandifer, 2003) but also suggests that designing for multiple visualizations within a mini-exhibit setup may garner deeper knowledge understanding.

Limitations

While these results strongly demonstrate evidence of increased learning with multiple visualizations, we recognize that this study is not without its limitations and challenges, some of which admittedly will have impacted the positive findings. The first limitation concerns the design of the research study. While there are several benefits to conducting a within-subjects study, one disadvantage is the potential for order effects to negatively impact the data. In this study, children were asked, "How do you think the ball is able to float without being blown away?" 3 times. The successive repetition of this question may have cued participants to pay more attention to the information in visualizations than if they had encountered them on the museum floor. Second, the study occurred in a location separate from the main museum floor – half in one of the design studio and the other half in an office-like space. These environments greatly contrast a typical museum environment where the scene is often chaotic, loud, and distracting with multiple activities going on at the same time. However, we view this as a "first study" and best-case scenario of what could occur. Families had the entire exhibit to themselves, the environment was void of typical museum distractions, and they could spend as long as they wanted. To understand how this mini exhibit would fare on the actual museum floor would require further research. Finally, some may argue that increased learning resulted more from the repetition of seeing 3 visualizations rather than from the actual information presented by the visualizations themselves. We actually consider this to be an integral advantage of having multiple visualizations. Several children commented in their interviews that it was helpful to have this reiteration as it lent "more proof and understanding rather than just one" and that "since you have all 3 that are kind of similar, you can review over and over again which inputted *[sic]* in your head." Simply having multiple visualizations served to reinforce a particularly difficult concept that children and adults alike, struggle with. Thus, the presence of redundant information should not be considered a casual accessory to be minimized but instead, as an asset that can improve learning.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, 33, 131-152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16,* 183-198.

Allen, S. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education, 88*(Suppl. 1), S17-S33.

Asai, K., Sugimoto, Y., & Billinghurst, M. (2010). Exhibition of lunar surface navigation system facilitating collaboration between children and parents in science museum. *Proceedings of the 9th ACM SIGGRAPH Conference on Virtual Reality Continuum and Its Application in Industry*, 119-124.

Basca, B., & Grotzer, T. (April, 2001). Focusing on the nature of causality in a unit on pressure: How does it affect student understanding? Paper presented at the *Annual Meeting of the American Educational Research Association,* Seattle, Washington.

Borun, M., & Dristas, J. (1997). Developing family-friendly exhibits. *Curator, 40*(3), 178-196.

Cheng, M., Annetta, L., Folta, E., & Holmes, S. (2011). Drugs and the brain: Learning the impact of methamphetamine abuse on the brain through a virtual brain exhibit in the museum. *International Journal of Science Education, 33*(2), 299-319.

de Jong, T. (2011). Instruction based on computer simulations. In R. Mayer & P. Alexander (Eds.), *Handbook of Research on Learning and Instruction* (pp. 446-466). NY: Taylor & Francis.

Eberbach, C., & Crowley, K. (2005). From living to virtual: Learning from museum objects. *Curator, 48*(3), 317-338.

Engel Clough, E., & Driver, R. (1985). What do children understand about pressure in fluids? *Research in Science and Technological Education, 3*(2), 133-144.

Falk, J. (1997). Testing a museum exhibition design assumption: Effect of explicit labeling of exhibit clusters on visitor concept development. *Science Education, 81,* 679-687.

Falk, J., & Storksdieck, M. (2005). Using the contextual model of learning to understand visitor learning from a science center exhibition. *Science Education, 89,* 744-778.

Hall, T. & Bannon, L. (2006). Designing ubiquitous computing to enhance children's learning in museums. *Journal of Computer Assisted Learning*, *22*(4), 231-243.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14,* 343-351.

Laursen, D. (2013). Co-participation among school children around a computer-based exhibit. *Social Studies of Science, 43*(1), 97-117.

Lowe, R. K., & Schnotz, W. (Eds.). (2008). *Learning with animation: Research implications for design*. New York: Cambridge University Press.

Matuk, C., & Uttal, D. (2008). Entertaining evolution: Understanding science from animations. *Proceedings of the 8th International Conference on the International Conference for Learning Sciences 3,* 93-94.

National Research Council (NRC). (2009). *Learning science in informal environments: People, places, and pursuits.* Washington, DC: National Academies Press.

Plass, J., Homer, B., & Hayward, E. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education, 21*, 31-61.

Rogers, Y. (2008). A comparison of how animation has been used to support formal, informal, and playful learning. In R. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (pp. 286-303). NY: Cambridge University Press.

Sandifer, C. (2003). Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a science museum. *Journal of Research in Science Teaching, 40*(2), 121-137.

Scheiter, K., Wiebe, E., & Holsanova, J. (2009). Theoretical and instructional aspects of learning with visualizations. In R. Zheng (Ed.), Cognitive effects of multimedia learning (pp. 67-88). Hershey, PA: IGI Global.

Schnotz, W., & Lowe, R. (2008). A unified view of learning from animated and static graphics. In R. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 304-356). New York: Cambridge University Press.

Sere, M. (1982). A study of some frameworks used by pupils aged 11-13 years in the interpretation of air pressure. *European Journal of Science Education, 4*(3), 299-309.

Ucko, D., & Ellenbogen, K. (2008). Impact of technology on informal science learning. In D. W. Sunal, E. Wright, and C. Sundberg (Eds.), *The impact of the laboratory and technology on learning and teaching science K-16* (pp. 239-266). Charlotte, NC: Information Age.

Van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning & Instruction,16,* 199-212.

Wang, J., & Yoon, S. (2013). Scaffolding visitors' learning through labels. *Journal of Museum Education, 38*(3), 320-332.

Yoon, S., Elinich, K., Wang, J., Van Schooneveld, J., & Anderson, E. (2013b). Scaffolding informal learning in science museums: How much is too much? *Science Education, 97*(6), 848-877.

Yoon, S., & Wang, J. (2014). Making the invisible visible in science museums through augmented reality devices. *Tech Trends*, *58*(1), 49-55.

Yoon, S., Elinich, K., Wang, J., Steinmeier, C., & Tucker, S. (2012a). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*, *7*(4), 519-541. doi:10.1007/s11412-012-9156-x

Yoon, S., Elinich, K., Wang, J., Steinmeier, C., & Van Schooneveld, J. (2012b). Learning impacts of a digital augmentation in a science museum. *Visitor Studies, 15*(2), 157-170. doi:10.1080/10645578.2012.715007.

The Role of Inconsistencies in Collaborative Knowledge Construction

Martina Bientzle, Ulrike Cress, Joachim Kimmerle, Knowledge Media Research Center,
Knowledge Construction Lab, Schleichstr. 6, 72076 Tuebingen
Email: m.bientzle@iwm-kmrc.de, u.cress@iwm-kmrc.de, j.kimmerle@iwm-kmrc.de

Abstract: When learners use shared digital artifacts for purposes of knowledge construction, they may be confronted with different types of inconsistencies. We present a study in which physiotherapy students had to handle two types of inconsistencies: Inconsistencies within a wiki text and inconsistencies between the text and their own attitude toward health and therapy. We conducted both a quantitative and a qualitative analysis of the students' contributions. We found that students modified inconsistencies within a text by changing the text, sometimes even by changing scientific facts. When learners revised inconsistencies between the text and their own attitude they did not change their personal therapeutic health concept, but transformed the text to fit their own perspective. Finally, we found that students rearranged scientific facts to a higher degree if the text was not in line with their attitudes. We discuss the implications of our findings for understanding knowledge construction.

Introduction

Shared digital artifacts play an increasing role in computer-supported collaborative learning and collaborative knowledge construction. This applies both to environments that are explicitly intended for purposes of knowledge building (such as Knowledge Forum; see Fujita, 2013; Hong & Scardamalia, 2014; Zhang, Scardamalia, Reeve, & Messina, 2009) as well as to tools that allow for a rather casual and incidental way of learning and knowledge construction (such as social tagging tools, for example; Cress, Held, & Kimmerle, 2013; Golder & Huberman, 2006). Shared artifacts allow their users to make individual contributions and to introduce their own knowledge to the community of all users. But different types of shared digital artifacts vary in the degree in which they can be influenced and modified. Tools that are particularly convenient for users to make comprehensive modifications are wikis (Kimmerle, Cress, & Held, 2010). In wikis users have the opportunity to drastically influence the content of the shared artifact (Leuf & Cunningham, 2001; Moskaliuk & Kimmerle, 2009).

When people use shared digital artifacts for purposes of knowledge construction, they are confronted with different types of inconsistencies. There may be inconsistencies within a shared artifact itself. This is the case, for example, when contradictory words or definitions are used to refer to the same object in a social tagging environment or when different statements contradict each other within a wiki article (e.g., Kimmerle, Cress, Held, & Moskaliuk, 2010). Inconsistencies within texts and their implications for cognitive information processing have been examined in text comprehension research for more than two decades (see Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Kintsch, Songer, & Kintsch, 1996). McNamara et al. (1996) argued, for example, that a poorly composed text may support skilled readers in elaborating on its content, since they have to cognitively compensate for imprecise or ambiguous information. When using wikis for knowledge construction, however, the collaborative handling of inconsistencies within the shared text should also be given attention, in addition to the individual cognitive processing of information (see Cress & Kimmerle, 2007, 2008; Kimmerle, Moskaliuk, & Cress, 2011), in particular since the selective handling of information may have implications for the collective construction of knowledge. So far there has not been enough investigation into how such inconsistencies affect the collaborative production of a shared text. Another type of inconsistency may occur between the knowledge, beliefs, or attitude of an individual user and the information contained in the artifact. It is an empirically well-supported finding that such contradictions may lead to some kind of cognitive dissonance (Festinger, 1964) or to socio-cognitive conflicts (Berlyne, 1960; Piaget, 1977). When this occurs, individuals then have to deal with the contradictions in order to reduce the dissonance and re-equilibrate their cognitive structures. If learning scientists want to implement collaborative learning settings that support an unbiased handling of information, they need to understand learners' strategies of dealing with contradicting attitudes. A particularly relevant question for the learning sciences is how people deal with these inconsistencies in situations where they are not limited to a purely cognitive processing of the inconsistencies. What happens when they have the opportunity to modify the content of a shared digital artifact—as in the case of wikis? How strongly are they bothered by inconsistencies within a wiki text? How do they act or react to content that does not suit their view of the world? And what is the impact of particular combinations of inconsistencies? That is, do users react differently to inconsistencies within a wiki text when its general message is in line with their attitude, compared to a text that contradicts their opinion?

Bientzle, Cress and Kimmerle (2013) reported that inconsistencies within a wiki text per se did not provoke wiki users to conduct more modifications than in a situation without such inconsistencies in the text. But in texts that expressed a point of view contradictory to that of the users, those with inconsistencies were modified more frequently than texts without inconsistencies. It is apparent that people tolerated logical inconsistencies to a much lesser degree when they appeared in the wrapping of a contradictory point of view. This selective handling of information in the collaborative construction of knowledge is relevant for the learning sciences. If researchers aim to support learners in contributing to a preferably open-minded and unbiased knowledge construction process, then they need to comprehend how learners act when they have to deal with inconsistencies. So it is worthwhile to take a much closer look at how the findings of Bientzle et al. (2013) came about. Previous literature has not sufficiently described the underlying processes and the strategies that learners apply in such situations. Therefore, we provide in this article a detailed qualitative analysis of people's contributions. First, we describe the procedure of the experiment. Then we present a variety of exemplary findings of how the participants dealt with various constellations of inconsistencies. Finally, we discuss the implications of our results for future research in learning and education.

Method

In order to better understand how learners handle different types of inconsistencies, we analyzed their contributions, deletions, and modifications in a wiki text. On the basis of Bientzle et al.'s (2013) data set, we examined under which conditions learners applied particular strategies for dealing with inconsistencies in a wiki text. Since medical information is particularly concerned with inconsistencies (Kienhues, Stadtler, & Bromme, 2011; Kravitz, Duan, & Braslow, 2004; Sniderman & Furberg, 2009), a health-related topic was chosen for this study, and physiotherapy students were recruited as participants.

Inconsistencies were incorporated into the text by adding logical contradictions into statements about the effectiveness of stretching. To establish inconsistencies between the text and a learner's attitude, we used a pre-test to measure the therapeutic health concept of the participants. Based on this measurement, we then provided a text that was either consistent with or contradictory to the therapeutic health concept of the participants. We constructed four different versions of texts by combining these two types of inconsistencies.

Participants

Seventy-six students (M_{age} = 21.99, SD = 4.28; men: n = 24, women: n = 50, 2 did not indicate their gender) of a state-approved school of physiotherapy (PT Academy) participated in the study. The students were in the first, second, or third (final) year of their vocational training. They were randomly assigned to one of the four versions of the text, which they edited individually at a laptop in a classroom setting.

Inconsistencies within a Text

The fluctuating quality of health knowledge and inconsistencies in health information are challenges for health-related knowledge construction. Health information is not only based on scientific facts, but very frequently also on practical and anecdotal knowledge. In many cases, however, anecdotal knowledge is not supported by scientific evidence at all. As a consequence, inconsistencies within health knowledge occur.

In the health care sector, especially in sports and physiotherapy, stretching is an established and widely used treatment. It is used to improve mobility and flexibility, to prevent injuries and aching muscles, and to foster a quicker regeneration. At the same time many scientific investigations disclose inconsistent and controversial effects of stretching (Small, Mc Naughton, & Matthews, 2008; Taylor, Dalton, Seaber, & Garrett, 1990; Weppler & Magnusson, 2010). Thus, this topic seemed to be particularly suitable to investigate what happens when students are confronted with relevant but inconsistent information in a situation where they have the opportunity to modify the content.

Inconsistencies between a Text and a Learner's Attitude

To investigate the effect of inconsistencies between content in a shared artifact and people's attitude, we chose to use people's therapeutic health concept. This is a central and well examined concept in the health sciences (Alonso, 2004; Laffrey, 1986). There are various perspectives on what health is considered to be. It is known that health professionals and patients differ in their individual understanding of health (Patel, Arocha, & Kushniruk, 2002), and that even among healthcare professionals no common understanding of the concept exists (Engel, 1977; Larson, 1999; Roberts, 1994).

In the health care sector two different classification systems coexist. The International Classification of Diseases (ICD; World Health Organization, 1992) promotes a *biomedical therapeutic health concept*, whereas the International Classification of Functioning, Disability and Health (ICF; World Health Organization, 2001) promotes a *biopsychosocial perspective*. The biomedical therapeutic health concept implies a scientific perspective on therapy and health that can be described by keywords like 'scientific', 'evidence-based', or 'standardized'. The biopsychosocial therapeutic health concept implies a more holistic perspective that can be

described by keywords like 'individualized', 'social participation', or 'functioning in everyday life'. In order to evoke inconsistencies between the text and a learner's attitude, we presented texts that were either consistent with participants' health concept or that contradicted their attitude toward therapeutic principles.

Experimental Material

By systematically combining inconsistencies, four different versions of the text were constructed: (1) text with consistent information and a therapeutic health concept that was congruent with the learners' health concept; (2) text with inconsistent information and a congruent health concept; (3) text with consistent information and a health concept that contradicted that of the learners; and (4) text with inconsistent information and a contradicting therapeutic health concept.

A text with inconsistent information included five statements that illogically connected a *scientific fact* with practical or *anecdotal knowledge*. An example of an inconsistent statement was: *"Temporary stretching resulted in no structural extension of the muscle fiber. This seems to be a reason for a better mobility after a stretching intervention."* In physiotherapy, stretching is widely used to increase the range of motion (ROM), in particular if muscle contractures are identified to be caused by a restriction of mobility. It is frequently assumed that a structural extension of the muscle fiber occurs via stretching, which is then supposed to be the reason for the positive effect on mobility. But the anatomic fact is that no structural extension of the muscle fiber occurs. This scientific finding diametrically opposes the anecdotal knowledge that an increase of the ROM would be due to the structural length of a muscle fiber. The discrepancy between these two statements is obvious to students of physiotherapy (manipulation check measures indicated that students estimated such text sections as more inconsistent; for details see Bientzle et al., 2013).

The therapeutic health concept of the text was operationalized by including either five statements congruent with the student's concept or five statements contradictory to the student's concept. As already known from the literature, the biopsychosocial perspective is common and popular in physiotherapy (Jorgensen, 2000; Stenmar & Nordholm, 1994). According to the pre-test of the health concept, this was also true of the participating sample of physiotherapy students. To take this into account, the text with a contradicting health concept included five biomedical statements, whereas the text with a congruent health concept included five biopsychosocial statements. An example of a contradicting (i.e., biomedical) statement was: *"Since stretching has various effects, patients benefit most from standardized examination and the use of evidence-based treatments."* An example of a congruent (i.e., biopsychosocial) statement was: *"Since stretching has various effects, patients benefit most from individualized examination and the use of treatments suited to the individual's needs."*

We conducted a qualitative analysis (Mayring, 2000) and a statistical frequency analysis in order to gain deeper insights into how students handle inconsistencies when they have the opportunity to modify the content of the text. We focused on those text sections that provided logical inconsistencies within the text and represented either a biomedical (contradicting) or a biopsychosocial (congruent) therapeutic health concept.

Results

In the following sub-sections we present the participants' strategies for handling inconsistencies within a text as well as inconsistencies between a text and learner's attitude (see Table 1 for an overview). We also discuss the effects of the interaction of these two types of inconsistencies.

Table 1: Strategies for handling inconsistencies.

Type of inconsistency	Strategy	Frequency (percentage)
Inconsistencies within a text	No modification	83 (46.7%)
	Deletion of the entire section of the text	32 (16.8%)
	Resolving the inconsistency reasonably	44 (23.2%)
	Modifying the scientific fact	31 (16.3%)
Inconsistencies between a text and learner's attitude	No modification	102 (53.7%)
	Revising without changing the perspective of the text	9 (4.7%)
	Deletion of the entire section of the text	21 (11.1%)
	Transforming into own perspective	28 (14.7%)
	Integrating contradicting attitude with the own attitude	30 (15.8%)

Inconsistencies within a Text

Over all conditions and participants, 190 sections represented inconsistencies within the text. Eighty-three (43.7%) sections were not modified at all. There were 11 students who modified five out of five inconsistencies and six students who never modified any inconsistencies. This quite large range of behavior indicates that the pure existence of logical inconsistencies does not guarantee active participation in knowledge construction in order to solve these discrepancies. In the following, we provide a deeper analysis of the remaining 107 (56.3%) text sections in which the students did modify logical inconsistencies.

Thirty-two (16.8%) text sections with logical inconsistencies were completely deleted. Students who deleted the entire section made an effort to create a coherent text by trying to get rid of these confusing segments. They were obviously able to identify those specific sections of the text that were a source of inconsistency. But they did not make the effort to look into the problem more deeply.

Forty-four (23.2%) logical inconsistencies were reasonably resolved with respect to content. In most of these cases students presented a solution to the inconsistency in that they deleted the inference that was incorrectly drawn from the scientific finding. This is illustrated by the modification made to the following example: *"It is already known that stretching fosters the range of motion. Temporary stretching resulted in no structural extension of the muscle fiber. ~~This seems to be a reason for a better mobility after a stretching intervention.~~"*

The selective deletion of an illogical inference needs—in contrast to deleting an entire section—a deeper understanding of the content. Students who performed such selective deletions had to be aware of the quality of information: It seems they considered scientifically evaluated information as more trustworthy than practical or anecdotal knowledge. They were able to identify the anecdotal knowledge as the source of inconsistency and accordingly drew the conclusion that it was this part that needed to be removed in order to achieve a coherent text.

Another approach to resolving an inconsistency was to adapt the anecdotal knowledge. The following example demonstrates this strategy. This was the original text: *"It is commonly considered that after temporary stretching the flexibility of the musculature is improved. A hint of this assumption is that the stress-strain curve of the muscle is left unchanged after stretching."* This inconsistency was adapted by simply inserting the word *"not"*: *"... that after temporary stretching the flexibility of the musculature is <u>not</u> improved. A hint of this assumption ..."*.

The adaptation of the anecdotal knowledge showed that students who applied this strategy apparently reflected about the inconsistency and were thus even able to incorporate the anecdotal knowledge in a modified way. It seems they were familiar with scientific reasoning to such a degree that they demonstrated the ability to interpret a scientific fact correctly.

An additional, even more demanding strategy to resolve an inconsistency reasonably was to adapt the anecdotal knowledge and add further information at the same time. The following text section is an illustrating example: *"Stretching is used to foster a quick regeneration after high physical load. The reduced blood flow and the traction on the stressed and micro-traumatized musculature during static stretching provide an explanation for the assumed positive effect of stretching on regeneration."* This inconsistency was adapted by inserting further information about missing scientific evidence. In addition, this student adapted the illegitimate conclusion from the scientific fact by deleting *"the assumed positive effect of stretching on"* and by inserting a correct conclusion. The whole section of the text appeared after the revision in the following way: *"Stretching is used to foster a quick regeneration after high physical load. <u>However, this has not yet been proved, since regeneration processes cannot be associated with stretching</u>. The reduced blood flow and the traction on the stressed and micro-traumatized musculature during a static stretching <u>do not</u> provide an explanation for ~~the assumed positive effect of stretching on~~ regeneration."*

Besides a deep understanding of the specific content, such an elaborate revision requires people's ability to reflect on the current stage of knowledge about the effectiveness of stretching. Students who were able to perform such revisions had to be aware of the uncertainty in medical anecdotal knowledge and also of the lack of scientific evidence in physiotherapy.

Apart from these strategies to resolve inconsistencies as described so far, inconsistencies within the text also stimulated the participants to modify the scientific fact. Thirty-one (16.3%) of the logical inconsistencies were edited in this way.

In the following example the scientific fact was adapted in order to fit to the anecdotal knowledge. This was the original sentence: *"The reduced blood flow and the traction on the stressed and micro-traumatized musculature during a static stretching ..."*. A participant changed this sentence in the following way: *"The ~~reduced~~ <u>enhanced</u> blood flow ..."*.

This is another example of a participant who revised the scientific fact: *"... the stress-strain curve of the muscle is left unchanged after stretching"* was changed into: *"... the stress-strain curve of the muscle is ~~left unchanged~~ <u>changed positively</u> after stretching."* After this student's revision the conclusion that the scientific

fact would be an explanation for the positive effect of stretching became logically correct (even though it was still incorrect with regard to current scientific knowledge).

Students who performed such adaptations had to be aware of the inconsistencies in the text. But in these cases, they identified the scientific knowledge as the source of the inconsistency. It seems they considered scientific knowledge as less certain and trustworthy than established anecdotal knowledge.

Inconsistencies between a Text and a Learner's Attitude

In total, 190 sections of the texts represented a biomedical therapeutic health concept which contradicted the participants' health concept. We found that 102 (53.7%) sections were not modified at all. This indicates that the mere existence of inconsistencies between a text and learner's attitude did not necessarily lead to text modifications. Perhaps students did not realize the inconsistencies or they were not personally affected by them.

In the following, we provide a deeper analysis of the remaining 88 (46.3%) text sections in which the students did modify the text. Nine (4.7%) biomedical statements were revised without changing the therapeutic health concept. In the following example only the extent of the standard ROM of the knee joint was modified: *"A deficit in the mobility which should be treated exists, if a patient does not achieve the standard ROM [i.e., knee: ext/flex ~~(0/0/180)~~ (0-5/0/120-150)]"*

It seems that students who performed such an adaptation were not bothered by the biomedical statement; they even tried to improve the accuracy of the biomedical information. This is remarkable in the face of the fact that all participants preferred biopsychosocial statements. This behavior seems to reflect some open-mindedness toward other attitudes or at least a willingness to tolerate other perspectives.

In contrast to this editing behavior we also found that 21 (11.1%) text sections that represented a biomedical therapeutic health concept were completely deleted. Deleting the whole section of the text was an option which had the effect of neutralizing attitudes represented in the text. Students who deleted the entire section made an effort to diminish the contradicting therapeutic health concept. It seems they were disturbed by the contradicting attitude but they did not invest much effort to transform the section of the text to fit their own perspective.

Another approach of some participants was to transform the contradicting statement to be in sync with their own perspective. Twenty-eight (14.7%) biomedical statements were transformed into biopsychosocial statements, whereas no biopsychosocial statement was transformed into a biomedical statement. The following example demonstrates this strategy. The statement *„Since stretching has various effects, patients benefit most from standardized examination and the use of evidence-based treatments"* was adapted by replacing the biomedical keywords *"standardized"* and *"evidence-based"* with the biopsychosocial keywords *"individualized"* and *"individual"* respectively.

Another approach to eliminating the biomedical perspective was to relativize the biomedical statement with the help of a biopsychosocial statement. For example, keywords like *"standard ROM"* remained unchanged. But at the same time the statement *"A deficit in the mobility which should be treated exists, if a patient does not achieve the standard ROM [i.e., knee: ext/flex (0/0/180)"* was revised by replacing the second part of the statement with the following change: *"... which should be treated exists, if a patient isn't able to handle his daily routine caused by the deficits. However treatment decision should not mainly depend on the standard ROM."*

Students who performed these strategies made an effort to diminish the contradicting therapeutic health concept and to adapt the text to their personal way of thinking. This behavior seems to reflect a very strong opinion and little willingness to accept other attitudes.

So far only strategies of acceptance or rejection have been presented, but there were also 30 (15.8%) statements for which students made an effort to integrate the contradicting attitude with their own attitude. The statement *"A deficit in the mobility which should be treated exists, if a patient does not achieve the standard ROM [i.e., knee: ext/flex (0/0/180)"* was supplemented by *"while the main focus should be on deficits that become apparent in the everyday life of the patient."*

Another student only replaced *"standardized"* with *"individualized"* and left the biomedical keyword *"evidence-based"* unchanged. The transformation of this biomedical statement appeared as follows: *"Since stretching has various effects, patients benefit most from ~~standardized~~ individualized examination and the use of evidence-based treatments."*

Students who exhibited such editing behavior made an effort to reach a balanced presentation of the therapeutic health concepts. They did not devalue the contradicting attitude. This seems to be a very sophisticated way of handling those inconsistencies between the opinion voiced in the text and their own attitude—particularly impressive since all students had a strong biopsychosocial orientation.

Interaction of the Different Types of Inconsistencies

A particularly interesting and key question is how the combinations of the two types of inconsistencies exerted influence on the contributions and the editing behavior of students. Therefore, we examined the students'

reactions to inconsistencies within a text when the text was in line with their personal attitude compared to the situation in which the text contradicted their biopsychosocial health concept. To test differences in the reactions to inconsistencies depending on the health concept of the text, we calculated chi-squared tests.

There were 95 sections that included inconsistencies within the text which were embedded in a health concept contradictory to that of the students. Another 95 sections contained inconsistencies within the text which were in line with the participants' health-related attitude. In the following, we present three types of revision strategies which were applied by the participants: Reasonable resolutions of inconsistencies, deletions, and rearrangements of scientific facts.

Overall, 61 (64.2%) text-immanent inconsistencies were edited in texts that expressed contradictory health concepts, while only 46 (48.4%) inconsistencies were revised in texts that were in line with the participants' personal attitude: $X^2 = 4.81$, $p = .028$.

This difference is not attributable to reasonable resolutions of inconsistencies, however. Twenty (21.1%) inconsistencies within biomedical texts and 24 (25.3%) inconsistencies within biopsychosocial texts were revised in this way: $X^2 = 0.47$, $p = .492$.

The differences regarding deletions were greater, though also not statistically significant. Here, 20 (21.1%) inconsistencies within biomedical texts and twelve (12.6%) inconsistencies within biopsychosocial texts were completely deleted: $X^2 = 2.41$, $p = .121$.

The strategy that contributed to the overall difference very strongly was the rearrangement of scientific facts. Here, the frequency differed significantly between conditions. Twenty-one (22.1%) inconsistencies within biomedical texts but only ten (10.5%) inconsistencies within biopsychosocial texts were revised by using this strategy: $X^2 = 4.66$, $p = .031$.

In sum, it seems that students were more tolerant toward logical inconsistencies within a wiki text when this text was in line with their personal attitude toward health and therapy. Our findings regarding deletions clearly support this interpretation. The attitude represented in the text, however, had no influence on the most sophisticated and challenging revision strategy, that is, to resolve the text-immanent inconsistencies reasonably. Finally, the cases where a biomedical (i.e., contradictory) text caused the participants to revise a scientific fact stand out. They give the impression that a contradictory attitude as expressed in a wiki text was even able to tempt the physiotherapy students to manipulate scientific content according to their personal preferences.

Discussion

When shared digital artifacts, such as wikis, are used as learning settings to foster collaborative knowledge construction, students are confronted with different kinds of inconsistencies. Because of the fluctuating quality of knowledge and its diverse sources, inconsistencies may frequently occur within a shared artifact. In the analysis presented here we examined how students handle different types of inconsistencies.

First, we considered how inconsistencies within a wiki text affected the collaborative production of knowledge. We found that inconsistencies within the text resulted in diverse reactions of the participants. Almost half of the inconsistencies were not modified at all. The remaining inconsistencies were either deleted, reasonably resolved, or resolved in that students changed scientific facts. It was rather surprising that the students tended to modify scientific facts to such an extent. This can possibly be attributed to the nature of their vocational training. As can be seen from the German training and examination regulations of physiotherapy (PhysTh-APrV, 1994), physiotherapy is more a practical than an academic type of training. This might be one reason why anecdotal knowledge was apparently considered to be more trustworthy. Another reason may be that the anecdotal knowledge as presented in our material was closely linked to the practical knowledge of the students. In this situation, practical knowledge seemed to be more robust than theoretical knowledge (see also Hascher, Cocard, & Moser, 2004).

In a second step, we considered how inconsistencies between a text and a learner's personal attitude affected the participants' contributions. We found that students also used various strategies to handle such inconsistencies. Again, about half of the inconsistencies were not modified at all. The remaining inconsistencies were deleted, revised without changing the therapeutic health concept, transformed to fit people's own perspectives, or integrated into their own point of view. That some students tried to transform biomedical statements into biopsychosocial statements strongly indicates that the therapeutic health concept is a very central and stable concept for physiotherapists' professional identity. That some students integrated the contradictory attitude with their own attitude might be caused by the social character of the shared digital artifact. As already reported in previous research (e.g., Kimmerle, Moskaliuk, Bientzle, Thiel, & Cress, 2012) people tend to revise shared artifacts in a rather deliberate and tactical way.

Finally, our qualitative analysis suggests that the co-occurrence of inconsistencies led to specific effects. We found that students questioned logical inconsistencies within the text to a lesser degree if the general text was in line with their personal attitudes. This means that students seemed to take a more critical look at text-immanent inconsistencies when the shared artifact contradicted their own attitude. This might be an

interesting starting point for researchers and practitioners who want to foster critical handling of uncertain information by learners. In addition, we found that a contradictory attitude in particular fostered the revision of scientific facts. Since we operationalized the inconsistencies between the wiki and the participants' attitude only through the biomedical vs. biopsychosocial therapeutic health concept, we cannot certainly conclude that this effect would also occur when other types of inconsistencies exist between a text and a learner's attitude. Whether this result is specifically due to the clash of biomedical and biopsychosocial health concepts or can be generalized to other inconsistencies is an open question for future research.

The selective handling of information has important implications for the collective construction of knowledge, and is highly relevant for the learning sciences. According to our qualitative analysis, it is important to be aware that the pertinent source of an inconsistency may exert an essential influence on how students deal with it. Learning scientists should take into account the complex interplay between individuals (their knowledge, beliefs, attitudes, and abilities) on the one hand and shared digital artifacts on the other. And they should also consider the interplay among different sources of inconsistencies, if they want to implement learning settings that support collective knowledge construction.

References

Alonso, Y. (2004). The biopsychosocial model in medical research: The evolution of the health concept over the last two decades. *Patient Education and Counseling*, *53*, 239-244.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* New York: McGraw-Hill.

Bientzle, M., Cress, U., & Kimmerle, J. (2013). How students deal with inconsistencies in health knowledge. *Medical Education, 47*, 683-690.

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers & Education, 60*, 59-73.

Cress, U., & Kimmerle, J. (2007). A theoretical framework of collaborative knowledge building with wikis: A systemic and cognitive perspective. In C. A. Chinn, G. Erkens, & S. Puntambekar. (Eds.), *Proceedings of the 7th Computer Supported Collaborative Learning Conference* (pp. 153–161). New Brunswick, NJ: International Society of the Learning Sciences.

Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning, 3*, 105–122.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, *196*, 129-136.

Festinger, L. (1964). *Conflict, decision, and dissonance* (Vol. 3). Stanford University Press.

Fujita, N. (2013). Online graduate education course using Knowledge Forum. In D. Suthers, K. Lund, C. Rosé, N. Law, & C. Teplovs (Eds.), *Productive multivocality in the analysis of group interactions* (pp. 375-381). New York: Springer.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science, 32*, 198-208.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*, 193-202.

Hascher, T., Cocard, Y., & Moser, P. (2004). Forget about theory—practice is all? Student teachers' learning in practicum. *Teachers and Teaching*, *10*, 623-637.

Hong, H. Y., & Scardamalia, M. (2014). Community knowledge assessment in a knowledge building environment. *Computers & Education, 71*, 279-288.

Jorgensen, P. (2000). Concepts of body and health in physiotherapy: The meaning of the social/cultural aspects of life. *Physiotherapy Theory and Practice*, *16*, 105-115.

Kienhues, D., Stadtler, M., & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the web: When expert information breeds laypersons' doubts about experts. *Learning and Instruction*, *21*, 193-204.

Kimmerle, J., Cress, U., & Held, C. (2010). The interplay between individual and collective knowledge: Technologies for organisational learning and knowledge building. *Knowledge Management Research & Practice, 8*, 33-44.

Kimmerle, J., Cress, U., Held, C., & Moskaliuk, J. (2010). Social software and knowledge building: Supporting co-evolution of individual and collective knowledge. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of the Learning Sciences* (Vol. 1, pp. 9–16). Chicago, IL: International Society of the Learning Sciences.

Kimmerle, J., Moskaliuk, J., Bientzle, M., Thiel, A., & Cress, U. (2012). Using controversies for knowledge construction: Thinking and writing about alternative medicine. *Proceedings of the 10th International Conference of the Learning Sciences* (pp. 1-8). Sydney: International Society of the Learning Sciences.

Kimmerle, J., Moskaliuk, J., & Cress, U. (2011). Using wikis for learning and knowledge building: Results of an experimental study. *Educational Technology & Society, 14* (4), 138-148

Kravitz, R. L., Duan, N., & Braslow, J. (2004). Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *The Milbank Quarterly, 82*, 661-687.
Laffrey, S. C. (1986). Development of a health conception scale. *Research in Nursing & Health*, *9*, 107-113.
Larson, J. S. (1999). The conceptualization of health. *Medical Care Research and Review*, *56*, 123-136.
Leuf, B., & Cunningham, W. (2001). *The wiki way. Quick collaboration on the web.* Boston: Addison-Wesley.
Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research, 1*(2), Article 20.
Moskaliuk, J., & Kimmerle, J. (2009). Using wikis for organizational learning: Functional and psycho-social principles. *Development and Learning in Organizations, 23*(4), 21-24.
McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and instruction*, *14*, 1-43.
Patel, V. L., Arocha, J. F., & Kushniruk, A. W. (2002). Patients' and physicians' understanding of health and biomedical concepts: Relationship to the design of EMR systems. *Journal of Biomedical Informatics*, *35*, 8-16.
Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. New York: Viking Press.
PhysTh-APrV (1994). Retrieved October 27, 2013 from Federal Ministry of Justice website: http://www.gesetze-im-internet.de/physth-aprv/BJNR378600994.html
Roberts, P. (1994). Theoretical models of physiotherapy. *Physiotherapy*, *80*, 361-366.
Small, K., Mc Naughton, L., & Matthews, M. (2008). A systematic review into the efficacy of static stretching as part of a warm-up for the prevention of exercise-related injury. *Research in Sports Medicine*, *16*, 213-231.
Sniderman, A. D., & Furberg, C. D. (2009). Why guideline-making requires reform. *JAMA: Journal of the American Medical Association*, *301*, 429-431.
Stenmar, L., & Nordholm, L. A. (1994). Swedish physical therapists' beliefs on what makes therapy work. *Physical Therapy*, *74*, 1034-1039.
Taylor, D. C., Dalton, J. D., Seaber, A. V., & Garrett, W. E. (1990). Viscoelastic properties of muscle-tendon units: The biomechanical effects of stretching. *The American Journal of Sports Medicine*, *18*, 300-309.
Weppler, C. H., & Magnusson, S. P. (2010). Increasing muscle extensibility: A matter of increasing length or modifying sensation? *Physical Therapy*, *90*, 438-449.
World Health Organization (1992). *International statistical classification of diseases and related health problems, tenth Revision (ICD-10)*. Geneva: World Health Organization.
World Health Organization (2001). *The international classification of functioning, disability and health (ICF).* Geneva: World Health Organization.
Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences*, *18*, 7-44.

Acknowledgments

Some of the quantitative findings (i.e., the statistical results of a coding and counting procedure) of this experiment were reported in Bientzle et al. (2013). The qualitative analysis as presented in the article at hand has not been published before.

More Than Just Plain Old Technology Adoption: Understanding Variations in Teachers' Use of an Online Planning Tool

Heather Leary, University of Colorado Boulder, 594 UCB, Boulder, CO 80309, heather.leary@colorado.edu,
Victor R. Lee, Mimi Recker, Utah State University, Logan, Utah 84322, victor.lee@usu.edu,
mimi.recker@usu.edu

Abstract: This paper examines variability in teachers' usage patterns as they interacted with an online teacher support tool, the Curriculum Customization Service (CCS), as part of their professional work. The CCS is a web application that supports teachers in planning, adapting, sequencing, and enacting differentiated instruction in Earth science education. By mining the usage log files of over 40 teachers who used the CCS over a yearlong period, we analyzed for variability using a framework developed in marketing research to characterize appropriation of technology. This analysis helped reveal different kinds of teachers' patterns along two dimensions: frequency and variability of use. We then turned to qualitative records of teachers' experiences during the year to better understand why those variations appeared. Focusing on the experiences of several teachers, we distilled "contextual contingencies" that influenced how they chose to appropriate and use the CCS.

Introduction

The complex work that constitutes teaching has and continues to comprise an important area of research in the Learning Sciences (Fishman & Davis, 2006). One function of that body of research (e.g., Horn, 2005; Schneider et al., 2005) has been to inform the design, implementation, and modification of new technologies to support various aspects of work that teachers do inside and outside of a typical school day. For example, technologies have been developed to support teachers as they learn how to use new, inquiry based curricula (Fishman, 2003), reflect on specific moments of their own teaching (Sherin & van Es, 2005), engage in assessment conversations with their colleagues (Shapiro & Wardrip, 2011), participate in online professional development (Dede et al., 2009; Schlager & Fusco, 2003), or seek ideas and suggestions from a distributed network of professional peers for their lesson planning (Renninger & Shumar, 2002; Recker et al., 2013; Sumner & CCS Team, 2010). Beyond the explicit decision to design such tools in a way to support specifically identified work practices associated with teaching, these tools will often capitalize upon emerging technical capabilities and developments that can make the tool more powerful and attractive for teachers to use. As such, technology tools designed for teachers bear similarity to many other technology products in that they tend to accrue a number of new features with each passing design iteration and version release.

This accrual of features means that while we can (and should) continue to be concerned with whether or not teachers will adopt a given technology, we also need to be mindful of *how* those who do adopt the technology will actually use it. To illustrate, consider the example of a smartphone. This ubiquitous device is now built with capabilities such as voice calling, text messaging, video chatting, schedule management, contact management, email, Internet access, and image capture. We would reasonably expect from our own experiences that the patterns of use associated with these varied features will differ across different groups of users. For example, many young adults use smartphones heavily in service of text messaging and image capture, but only occasionally for voice calls. Working professionals may use their smartphones largely for email, schedule management, and contact management in a manner quite different from a senior citizen who may prefer to use his/her smartphone for voice calls and video chats with friends and family. While we could consider all of these individuals to be "adopters" by virtue of owning a smartphone, their actual patterns of use will greatly differ.

In this paper, we hypothesize that similar variability in use patterns will also appear for teachers as they engage with new, feature-rich technology tools intended to support their professional work. Our primary goals with this paper are the acknowledgement and characterization of these use patterns, followed by identification of what we have termed "contextual contingencies" that lead to these use patterns. We argue that the outcome of this work will be profitable and necessary for the Learning Sciences community as it continues to design and improve technology innovations to better support teachers.

In the remainder of this paper, we provide the research context and theoretical framework for this work, followed by a description of the research design, data sources, and analysis. We then present examples of five teachers, through which we explore variability in usage patterns as teachers use an online planning tool, and analyze teachers' experiences to help understand how contingencies influenced variation in these usage patterns.

 © ISLS

Research Context

The Curriculum Customization Service (CCS) is a web-based application that supports teachers in planning, adapting, sequencing, and enacting differentiated instruction. Developed through a participatory and iterative design process with several practicing teachers, the tool integrates research-based inquiry-focused curricula (publisher materials, specifically the *EarthComm* curriculum published by *It's About Time*) with open educational resources (OERs) and teacher contributed materials in the context of learning goals and key curricular concepts (Sumner & CCS Team, 2010). The CCS currently includes content for middle and high school Earth science, middle school physical science, and most recently, high school algebra. The focus of the current study is high school Earth science teachers.

In response to teacher feedback and observations of actual CCS use by teachers, a number of features have been added to better support teachers' work practices (see Figure 1). For example, the CCS has features in it now that allow teachers to match online and publisher resources to district learning goals, store a personalized set of preferred resources for later access, and build custom sequences of instructional materials from discovered resources that they can access while teaching. Additionally, features associated with new social media platforms have also been added. In particular, teachers can assign and view star ratings and descriptive tags for resources, see the number of people who have stored a particular online resource, and display a live activity stream indicating recent and current usage of materials by other teachers.

To date, there have been a number of noteworthy successes with the use of the CCS in several schools (Butcher, Ferrara, & Devaul, 2013; Sumner & CCS Team, 2010; Ye et al., 2013). Teachers report that the CCS and its various features (that many of them had suggested) has helped them to approach their daily teaching in new ways, to become aware of the colleagues' work practices, and to customize their instruction for their students' specific needs. Student exam scores indicate that students, especially English language learners, are increasing in their knowledge and understanding of science concepts. Yet even with these successes, the increase in the number of features in the CCS led us to suspect that the ways in which teachers use the tool may vary in consequential ways.

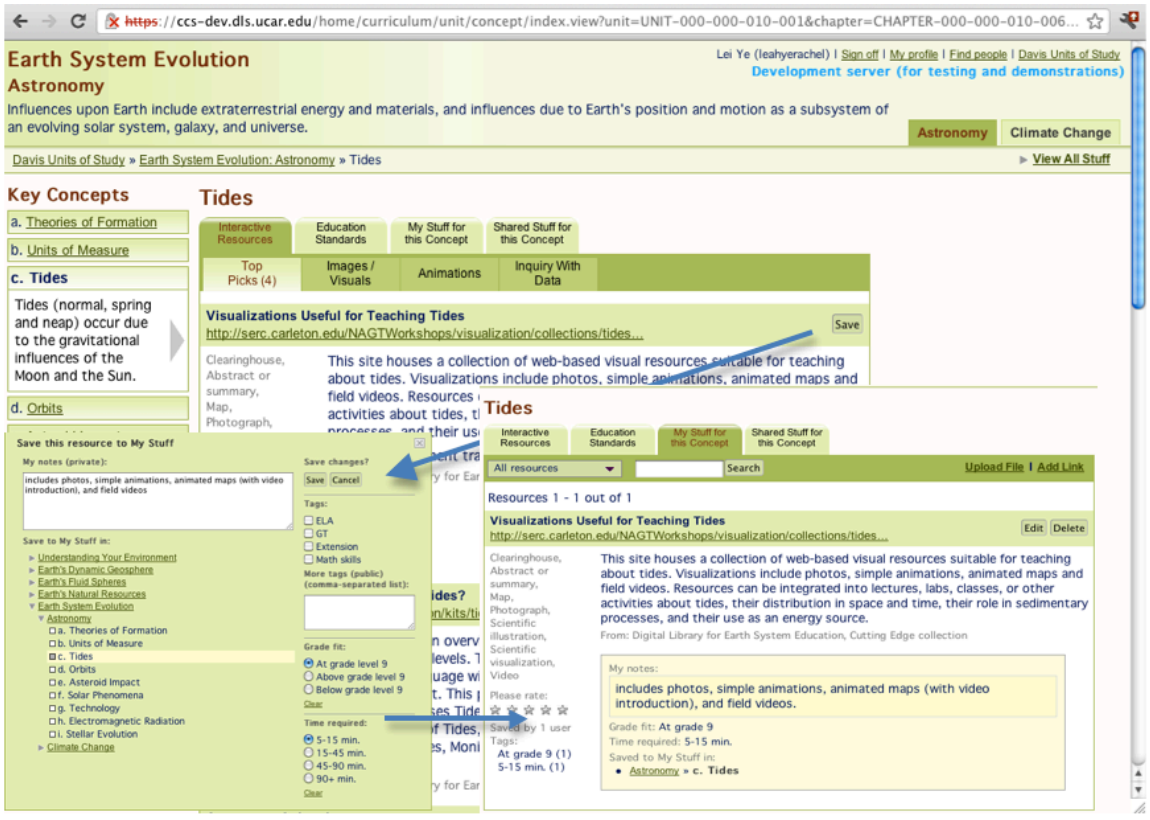


Figure 1. A screen capture of the CCS as a user selects an online resource and saves it as a private resource.

Theoretical Perspectives and Related Prior Work

This study of variation in teachers' use of the CCS is informed by two theoretical perspectives related to appropriation of new information technologies "in the wild." The first is a perspective on technology appropriation called "use diffusion" that originated from the marketing research literature (Shih & Venkatesh, 2004). Similar to the example of smartphones presented above, the use diffusion perspective also begins with

the assumption that different groups of people will preferentially engage with different features of a new technology. Indeed, this is where Shih & Venkatesh argue that progress is more likely to be made in efforts to better characterize consumer behavior. To capture the predicted variation, they propose modeling use in terms of two dimensions: frequency of technology use and variety of feature use (Figure 2). Against this space of possible technology behaviors, there are four expected use patterns. These include *intense use* (high frequency and high variety), *specialized use* (high frequency and low variety), *non-specialized use* (low frequency and high variety), and *limited use* (low frequency and low variety). To our knowledge, use diffusion has only minimally been used as an approach to conceptualizing appropriation of technologies within educational settings (Maull, et al., 2011; Pennington, 2004), as much of the existing discourse has often focused simply on technology adoption (e.g, Rogers, 2003). If our hypothesis that feature usage in the CCS is diffuse, then use diffusion theory may be a useful framework for capturing that usage pattern variation in practice.

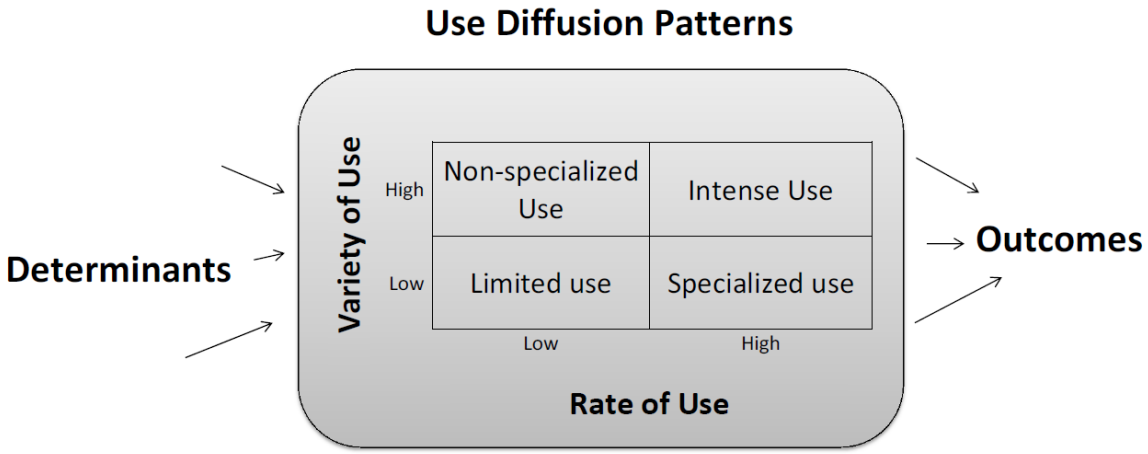


Figure 2. A depiction of Shih and Venkatesh's (2004) use diffusion model.

Yet, while use diffusion is helpful for characterizing how usage of a technology differs, it does not capture *why* those variations appear. Because of this limitation, we have also been drawing upon socio-technical systems perspectives to help us understand some of the ways that people shape and are shaped by their technology use (e.g., Bowker, Star, Tuner, & Gasser, 1997; Hollan, Hutchins, & Kirsh, 2000). These approaches typically focus on factors in a person's context that ultimately affect how technology is used, such as access to other tools or the norms of a local community. We refer to these mediating influencers collectively as "contextual contingencies."

As the CCS is relatively new, research on how the tool is used is still limited. However, one relevant example of recent work has highlighted how such contextual contingencies influenced use of the CCS within different school districts in a single academic year (Lee, Recker, & Sumner, 2013; Lee et al., in press). Specifically, the authors argued that the local culture with respect to sharing instructional resources and the policies of a district could both have influence on the degree to which the CCS was used by a given district. While that study successfully identified factors influencing teachers' use of the CCS, it was limited in that it only considered a single dimension of use (frequency). Also, it treated the school district as the primary unit of analysis, rather than the specific teachers who worked in those districts. When an analysis is focused at the level of individual teachers and on two dimensions of use, we anticipate additional variation. This variation we believe can also be attributed to more specific contextual contingencies.

Research Design and Methods

Our data come from a multi-site, yearlong study of over 70 9th grade Earth science teachers in 5 different school districts in the western United States. The districts were all provided with training resources to help them navigate the CCS and learn about its various features, and user accounts were provided for each teacher in the study so that they could discover, share, and recommend online resources to others in their school district.

Our analyses were guided by the following research questions. First, what does the use diffusion framework reveal about variability in teacher usage patterns in the CCS and, second, why do these variations occur?

Data Sources and Analyses

Throughout the yearlong period of research into CCS usage, we obtained both computational and qualitative data from the five school districts. Our computational data included automatically collected usage log files that

recorded each teacher's online activities (clickstreams) within the CCS environment. Automatically collected clickstream data have begun offering a number of insights into the behavior of students in technology-supported learning environments (Baker & Yacef, 2009; Bienkowski et al., 2012). While clickstream data from teachers have not been featured as prominently in the growing body of research on educational data mining and learning analytics, initial efforts to use clickstream data as recorded in other online tools to understand teachers' online behavior have been promising (Xu & Recker, 2012). The automatically collected CCS clickstreams included data about time, date, duration of login, types of resource accessed (e.g., publisher materials, OERs, etc), and operations performed on accessed resources such as whether it was opened, added, or removed from teachers' collections.

To analyze these data, we first completed a recommended phase of log file data cleaning and extraction, e.g., parsing log files, verifying accurate records, and performing some transformations on raw data, such as a logarithmic transformation on the number of CCS logins (Bienkowski et al., 2012). We then reduced our computational data to include only *active CCS users* from across the five school districts. As is often the case, there were teachers who expressed initial interest in the CCS but did not use it during the school year. We operationalized active users as those teachers who had logged in more than 3 times during the year (N=43). Next, to examine variety of usage, using a binary count, we coded the usage logs to see if a teacher had ever used a particular CCS feature (+1) or not. A use variety index score represented a sum of that count. We then used a best fit line between a plot of these two dimensions (the number of logins [log transformed] and use-variety score), and a line orthogonal to this best fit line to partition were used to partition teachers into the four quadrants recommended by Shih & Venkatesh (2004) (see Figure 3). This partitioning allowed us to identify, relative to our clickstream data, what we might consider *specialized use* rather than an *intense* or *limited use*.

Without a priori knowledge of what the usage patterns would be, we also recorded interview data from 26 participating teachers throughout the year. A minimum of two trained researchers jointly interviewed the 26 participants by phone at least twice during the school year using a semi-structured interview protocol designed to help the research team understand current district initiatives and pressures, in addition to individual teachers' perceptions and uses of the CCS. Also, 15 of those teachers were observed at least one day within their classroom settings on days when they intended to use the CCS so that we could see and record firsthand how the CCS was being used.

We then developed a grounded coding scheme of recurring themes in the interview data that was iteratively refined until it was systematically applied to the entire corpus and then cross-validated by three analysts across the entire interview corpus. Following that coding, we compiled dossiers of each teacher for whom qualitative data had been collected and then used the entire range of data at our disposal (e.g., interviews, clickstreams, observations, etc.) to prepare brief case reports of the teachers. After the teachers were mapped onto the use-diffusion matrix, we then examined specific examples and sought to identify from our case reports the contextual contingencies that could help to explain why a teacher's pattern of CCS use throughout the year unfolded the way that it did.

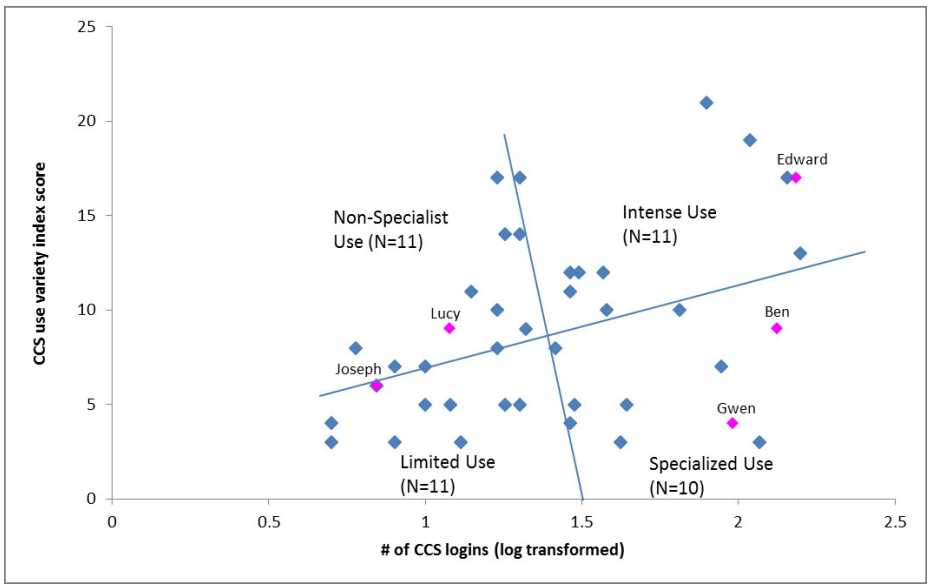


Figure 3. CCS active teacher users in the use diffusion framework and the selected teacher examples.

Results

With respect to our first research question, Figure 3 indeed shows variations along the two dimensions of frequency and variety of use. Among active users, the use variety index ranged from 3 to 21, and the number of (untransformed) logins ranged form 4 to 157. Depending on the type of usage (e.g., specialized use, intense use), the features used most included accessing digital versions of the textbook and embedded assessments from the publisher materials, along with seeking images and visualizations from OERs. The least used features were teaching tips in the publisher materials and the sets of inquiry data provided in OERs. Accessing shared materials showed a wide range of use. This approach to looking at teacher usage was especially helpful as teachers rarely reported on the specifics of what they accessed. However, even though this approach for looking at clickstream data helped us to see that variation along both dimensions, it did not tell us *why* this variation appeared. Thus, to address the second research question, we present examples of teachers from each of the quadrants (see Figure 3), and, when relevant, the contextual contingencies at play that could explain some of this variation.

Example 1: Intense Use Fitting with an Intended Use

Edward (all names are pseudonyms) was a teacher with eight years of teaching experience and was the lone Earth science teacher at his high school in a school district that had made a very strong commitment to using the adopted publisher materials as their curriculum of choice (as reflected by the financial investment, consistent use of publisher materials reported by other teachers in his district, and custom professional development activities led by the publisher of the materials). His contextual contingencies made him well-suited as a user of the CCS because he did not have immediate access to specific content area colleagues at his school who could help him in his topic specific planning or resource gathering routines. In fact, Edward had the highest number of logins in his school district, which already had a very high level of CCS usage relative to all the other districts in the study. Perhaps unsurprisingly, partitioning of teachers in terms of the use diffusion framework placed Edward into the *intense user* quadrant. Indeed, his reported CCS activity patterns fit quite well with how the CCS was originally designed to be used. When asked about whether and how he used the CCS, he responded emphatically that he did use the CCS and that it was beneficial for him because:

> All of the resources that are provided from [publisher materials] are there…I use it [the CCS] for storing things that I find [through repositories] that are useful and I want to include with all of my resources. I can save things to it and keep it in the [learning goals] section where it's useful.

Even from that brief transcript excerpt, we can catch a glimpse into the variety of Edward's CCS use. Edward was essentially reporting that he was accessing the embedded materials provided by the publisher of the adopted textbook, that he was using the CCS as a storage tool to help him remember what resources he had identified and thought would be useful, that it was a portal for him to explore OERs above and beyond what was provided by the textbook publisher, and that it was an organization tool for matching resources to specific lessons and learning goals.

Beyond the above transcript excerpt, Edward also reported that he was using the CCS to find resources that had been marked in the system as recommended by other teachers in his district. Additionally, he also reported that the CCS was an aid in his efforts to differentiate his instruction. For instance, he reported using the CCS to find resources he could download and modify for use with the subset of students who were English language learners because he felt that "they need[ed] more attention for reading and reading skill development."

Edward also separately indicated that use of the CCS led to changes in his lesson planning routine. As he put it, "I don't use a regular [paper based] plan book anymore. I can just plan it all online and then pull it up on my computer [during class]. That's a lot handier for me."

In many respects, this was the kind of use case for which the CCS was designed. Because he was the sole Earth Science teacher at his school, Edward was relatively isolated from his colleagues. He also had a diverse enough student population that required him to do some additional work to curate a set of resources that could be customized for his classes. These factors all worked in such a way that he was enabled and implicitly encouraged to become an *intense user*.

Example 2: Non-Specialized Use Due to Preferred Alternatives

Lucy was in her third year of teaching and also served as the lead teacher for Earth science at her high school. She worked in a different school district that was much larger than Edward's. Through various interviews in this school district, we discovered that Lucy and her colleagues were facing a number of unique local community pressures that ultimately resulted in the CCS having a diminished priority within new district-level reform initiatives. As a result, her school district had relatively low overall usage of CCS relative to the other school districts.

When we spoke with Lucy in her interviews, she reported a genuine interest and enthusiasm for online resources for teaching, but she did not rely on the CCS in the same way that Edward did. She treated the various resources as all being generic things to find on the Internet:

> As a teacher I use the Internet a lot to find sources that will help, you know guide whatever I'm going to be doing in class. I do it for my own knowledge of the material before I present it to the kids to make sure, you know I can kind of back it up too and give them a different source to look at it if they want to expand their learning.

When asked to elaborate on the online resources she used, she reported visiting a number of sites and services such as *Edmodo*, streaming video services, and even simply running Google searches. After some more direct questioning about the CCS, Lucy did report using and appreciating the CCS as one of many tools, but that did not come up until she was asked about the CCS specifically. When she did talk about the CCS, she commented that it was especially useful for helping her find specific animations and simulations to use in class:

> It's [the CCS] been really helpful to find those animations and to find those simulations, especially when you're teaching a class. For instance, when I was teaching waves there were kids who weren't understanding what I was trying to explain to them. I couldn't think of how to explain it differently. Originally my first thought was to pull up Google and try to do a quick Google search and then I remembered that there is a category on the CCS that has the animations and I went through it there and immediately I found the simulation that they use for the *slinkies*. And [that was] exactly what I was trying to use.

While Lucy was speaking in a manner complementary of the CCS, it is worth noting that her "first thought" was to "do a quick Google search" until she "remembered that there is a category on the CCS that has the animations". Thus, the CCS was not her primary source in the same way as it was for Edward. Like many other people, she turned to a reliable and widely known search engine that already serves a number of information discovery purposes. Lucy's pattern of using a broad range of technology tools in support of her teaching seemed to work for her in that while she may not have used the CCS as her primary platform for her teaching, her prior exploration of many different features as a *non-specialized* user of the CCS provided her with enough familiarity that she could remember there was an animations category that could, on occasion, serve her purposes. For Lucy, the contextual contingencies at play appeared to be a lack of a strong district encouragement to use the CCS and what appeared to be a combination of a comfortable familiarity with and access to a range of tools that could provide a number of other resources sufficient for her teaching needs.

Example 3: Specialized Use for Certain Features

Ben and Gwen were both *Specialized users* in terms of their frequency and patterns of use of the CCS, but interview data revealed that their specialized use patterns of the CCS were for very different reasons. Ben was in his first year of teaching high school Earth science, and although he was very comfortable using technology, he was less confident in the content he is teaching. He used two areas of the CCS, the materials shared by other teachers (referred to by a link called "shared stuff") and the publisher provided materials. He reported to us in his first interview that, "it's really been helpful to look at the "shared stuff", things other teachers have put together, good websites they have found for interactives and activities and whatnot." During his second interview, he reiterated his appreciation for this area of the CCS. He also mentioned being less familiar with the content he was teaching, which was why he used the shared area and publisher materials:

> "Being my first time teaching the content and obviously first time using this textbook and this curriculum, it was great to get on there [the CCS] the first couple of week and see how other teachers had done something or the resources that they were using. It's been great to collaborate with other teachers and wonderful networking."

Gwen, on the other hand, typically used the CCS to access only the publisher provided materials, and specifically PDF versions of textbook pages. She was a more seasoned teacher, in her eighth year of teaching Earth science. Her student population included a high number of English language learners. She frequently logged into the CCS and then logged out, with her average time on the site being less than four minutes according to the clickstream data. Her interviews revealed the reasons for her usage patterns: "I access the CCS to access the textbook so I can project it onto my screen." She stated that projecting the textbook provides all students (especially those who have forgotten their book) the opportunity to follow along and "no excuse for not having a textbook." She also had students read aloud from the projected PDFs, which she thought was useful for her many English language learners. In sum, both *Specialized users* showed high frequency of use with the CCS

but low variety use of the tools' features, reflected in their different teaching needs and preference for feature use.

Example 4: Limited Use Due to Technology Skills and Other Options

Joseph fell into the *Limited use* category. He had taught Earth science for four years, had abundant access to technology (he taught in a school with a 1:1 student:laptop ratio), and he was very technology savvy. In his interviews he reported the various projects he was working on within his school and district, including his use of Google Docs and the Prezi zoomable presentation service. He and a colleague were also working on creating an interactive digital curriculum map for the district, mapping the state curriculum standards using Prezi.

> "So the Prezi will- [when] you go into it, it outlines an inquiry lab for each, um, standard, objective I guess each objective in each standard, and then it sends it out to resources within Google Docs. You know, any sort of worksheet or lab write-up they'll need will be in shared folders within Docs. So it's kind of a combination of the two, the way we use it."

During the school year, his school district had actually not yet adopted a textbook, so no publisher materials were available in the CCS for teachers in this district. He did have access to all of the open educational resources and a district-wide shared resource collection. At the beginning of the school year Joseph commented to us that he was exploring the CCS and seeing "there wasn't a lot of [online] resources for the topic I happened to be on." He tried again at different points throughout the year but was again not satisfied with what he found. He stated, "I have a pile of other teachers materials I already use and I just go to Google and type in what I want and can usually find really good resources that way." Teachers in Joseph's district were in their first year of using the CCS, and from analysis of the district, we found that all the teachers had many other alternative means to share resources with one another. Consequently, little was added to the shared resources area in the CCS. Joseph indicated he was excited about the CCS and would have been more likely to use the CCS along with the long list of other online resources and tools he typically used if the shared area had contained more materials. However, the district context in which he worked, coupled with his familiarity with other tools, diminished the need he perceived to use the CCS or any of its features.

Discussion and Conclusion

In this paper, we examined variability in usage patterns as teachers used an online tool designed to support instructional planning, the Curriculum Customization Service (CCS). Like recent research examining student learning in online environments, we looked to usage log files to extract teacher patterns, and then analyzed these using the use diffusion framework. This analysis helped reveal different patterns of CCS use along the dimensions of frequency and variability of feature use. We then examined qualitative records of teachers' experiences during a single academic year, principally teacher interviews and classroom observations. Focusing on the experiences of teachers from each of the four use diffusion quadrants, we distilled some of the contextual contingencies that influenced how they ultimately used the CCS.

This approach demonstrates the potential for understanding teacher's online behaviors as a source of meaningful patterns of automatically collected usage data. In much of the recent discussion related to educational data mining and learning analytics (Bienkowski, et al. 2012), students' button clicks and online actions have been made the primary focus. But as many learning scientists are well aware, teachers play a critical role in structuring a formal learning experience. We feel that it is important that attention continue to be paid to the work of teaching, especially as new computational methodologies emerge. Moreover, we must continually anchor our interpretations of results from computational analysis with other known methods.

We believe that the approach modeled here with the CCS demonstrates that. It also may help encourage the field to recognize that teacher technology use more than simply a question about adoption. Rather, use of an "adopted" technology can vary greatly from one group of teachers to another. This variation can be made more visible when the use of different features is considered as an important dimension in how teachers work with technology. Of course, much work remains to be done to understand the range of factors that influence how a given support tool is appropriated by different teachers. However, our initial multi-method look at examples of teachers distributed across the space we have mapped out suggests there may still be much to learn at the level of an individual classroom and an individual teacher.

References

Baker, R. S. J. D., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Datamining, 1*(1), 3-17.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.* Washington, DC: Office of Educational Technology, US Department of Education, 1-57.

Bowker, G., Star, S.L., Turner, W., & Gasser, L. (Eds.) (*1997*). *Social science, technical systems and cooperative work: Beyond the great divide.* Hillsdale, NJ: Erlbaum

Butcher, K., Ferrara, L., & Devaul, H. (2013, June). Teachers' use of an online curriculum planning tool: Usage patterns associated with student learning. *Proceedings of EdMedia 2013: World Conference on Educational Media & Technology.* Chesapeake, VA: AACE.

Dede, C., Ketelhut, D. J., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A research agenda for online teacher professional development. *Journal of Teacher Education*, *60*(1), 8-19.

Fishman, B. J. (2003). Linking On-Line Video And Curriculum To Leverage Community Knowledge. *Advances in research on teaching*, *10*, 201-234.

Fishman, B., & Davis, E. A. (2006). Teacher learning research and the learning sciences. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (pp. 535-550). New York: Cambridge University Press.

Hollan, J. D., Hutchins, E. L., & Kirsh, D. (2000). Distributed cognition: A new theoretical foundation for human-computer interaction research. *ACM Transactions on computer-human interaction, 174-196*.

Horn, I. S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cognition and Instruction, 23*(2), 207-236.

Lee, V. R., Leary, H., Sellers, L. & Recker, M. (in press). The role of school district science coordinators in the district-wide appropriation of an online resource discovery and sharing tool for teachers. *Journal of Science Education and Technology*, doi: 10.1007/s10956-013-9465-5

Lee, V. R., Recker, M., & Sumner, T. (2013). *Variable appropriation of an online resource discovery and sharing tool*. Paper presented at the CSCW In Education Workshop, San Antonio, TX.

Maull, K., Saldivar, M. G., & Sumner, T. (2011). *Understanding digital library adoption: A use diffusion approach*. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries* (pp. 121-130), ACM, New York.

Pennington, M.C. (2004). Cycles of innovation in the adoption of information technology: a view for language teaching. *Computer Assisted Language Learning, 17*(1), 7-33.

Recker, M., Yuan, M. & Ye, L. (2013). *CrowdTeaching: Supporting Teachers as Designers in Collective Intelligence Communities*. Manuscript under review.

Renninger, K.A. & Shumar, W. (2002). Community building with and for teachers: The Math Forum as a resource for teacher professional development. In K.A. Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace*. New York, NY: Cambridge University Press.

Rogers, E. M. (2003). Diffusion of Innovations, 5th Edition. New York, NY: The Free Press.

Schlager, M. S., & Fusco, J. (2003). Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse?. *The Information Society*, *19*(3), 203-220.

Schneider, R. M., Krajcik, J., & Blumenfeld, P. (2005). Enacting reform‐based science materials: The range of teacher enactments in reform classrooms. *Journal of Research in Science Teaching*, *42*(3), 283-312.

Shapiro, R. B., & Wardrip, P. S. (2011). Interactive representations of student activity to inform teacher collaborations: Results from a formative exploration. In *Proceedings of CSCL2011 9th International Computer-Supported Collaborative Learning Conference* (pp. 494-501).

Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education*, *13*(3), 475-491.

Shih, C., & Venkatesh, A. (2004). Beyond adoption: Development and application of a Use-Diffusion model. *The Journal of Marketing, 68*(1), 59-72.

Sumner, T. & CCS Team. (2010). Customizing Science Instruction with Educational Digital Libraries. In *Proceedings of the 10th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2010)* (p. 4), New York: ACM.

Xu, B., & Recker, M. (2012) Teaching Analytics: A Clustering and Triangulation Study of Digital Library User Data. *Educational Technology & Society Journal, 15*(3), 103-115.

Ye, L., Walker, A., Recker, M., Leary, H., Devaul, H., Butcher, K., & Sumner, T. (2013, April). *Integrating Technology, Curriculum, and Online Resources: A Multilevel Model Study of Impacts on Science Teachers and Students*. Poster presented at the American Educational Research Association annual conference, San Francisco, CA.

Acknowledgments

We thank the teachers (and their students) who participated in our study. This work was supported by a grant from the National Science Foundation (DUE-1043858 and DUE-1043638). The opinions expressed herein are those of the authors and do not necessarily reflect those of the NSF.

Using an Adaptive Expertise Lens to Understand the Quality of Teachers' Classroom Implementation of Computer-Supported Reform Curricula in High School Science

Susan Yoon, Jessica Koehler, Joyce Wang, Emma Anderson, University of Pennsylvania
Email: yoonsa@gse.upenn.edu, jkoehl@gse.upenn.edu, joycew@gse.upenn.edu, ejanderso@gmail.com

Eric Klopfer, Massachusetts Institute of Technology, klopfer@mit.edu

Abstract: The exploratory study reported here is part of a larger-scale research project aimed at building theoretical and practical knowledge of complex systems in students and teachers with the goal of improving high school biology learning. In this paper we propose a model of adaptive expertise to better understand teachers' classroom practices. Through three case studies, we further illustrate the characteristics of adaptive expertise of more or less successful teaching and learning.

Introduction and Framing the Issue

The study of complex systems in the sciences and social sciences has become increasingly essential to understanding disciplinary and interdisciplinary content and practices (The National Academies, 2009). Likewise, in education, research on teaching and learning about complex systems has achieved solid grounding as an important field within learning sciences research (Hmelo-Silver & Kafai, 2011). In terms of educational policy and classroom enactment, the study of systems is also featured prominently in the *Next Generation Science Standards* (NGSS). Complex systems are generally defined as existing when any given numbers of interconnected and interdependent parts interact. The patterns of interactions form a network of relationships that exhibit emergent properties that cannot be observed or decomposed at subsystem levels.

Over the last couple of decades, learning sciences researchers have developed valuable resources and computational models for learning about complex systems such as StarLogo and NetLogo. However, few studies have focused on how best to scaffold the learning experiences for students in classrooms when the complex systems resources are expected to be integrated into standard science courses. Many studies instead have investigated the efficacy of those resources on learning (e.g., Yoon, 2008; 2011) without fully considering the contextual factors, including teacher differences that may be in play when implementing new resources. Because of the situated nature of teachers, classrooms, and schools, teachers' implementation of projects may diverge from developers' intentions in order to fit the learning needs of their contexts (Penuel et al., 2011). Hmelo-Silver & Azevedo (2006) have also pointed out that more research is needed on how to train teachers in complex systems resources and approaches to better support curricular and instructional experiences.

Given the importance of systems in the NGSS, the time is ripe to address these needs for developing teachers' knowledge and skills. A preliminary question to ask in terms of research design is what theoretical lens might be appropriate for understanding how to in-service teachers, and furthermore, for assessing the efficacy of teacher development. In one prominent line of research focused on understanding technology integration, researchers have identified barriers such as teacher beliefs, readiness, and a steep learning curve (e.g., Aldunate & Nussbaum, 2013; Ertmer et al., 2012). Others have discussed the importance of more exposure to computers (Mueller et al., 2008) or extensive computer training (Pierson et al., 2001). However, none of these studies offer information about or images of teachers who are in the process of becoming experts in computer-supported curricular integration. Thus, capturing how teachers enact reforms and evaluating their ability to adapt in order to identify levels of expertise can be instructive for teacher training purposes. We argue in this paper, that an adaptive expertise lens can offer support for understanding how teachers are able to navigate contextual factors (Penuel et al., 2011) while at the same time providing a framework to understand the quality of classroom implementation of computer-supported complex systems reform curricula in high school science. The exploratory study reported here is part of a larger-scale research project aimed at building theoretical and practical knowledge of complex systems in students and teachers with the goal of improving high school biology learning. In this paper we propose a model of adaptive expertise. Through three case studies, we illustrate the characteristics of adaptive expertise of more or less successful teaching and learning. By doing this research, it is our ultimate goal to contribute to scholarship on practices and training that teachers participate in to support complex systems teaching in classrooms.

Adaptive Expertise

Improving high school science teaching in the US has increasingly become a policy and practical imperative. High on the list of reasons are to meet the needs of a STEM-literate workforce (NRC, 2011); to produce better

 © ISLS

problem-solvers (OECD, 2011); and to support better decision-making about issues that can impact students' daily lives (Yoon, 2011). The number of reform-oriented STEM projects are far too numerous and varied to discuss here. However, the millions of dollars that the National Science Foundation alone pours into research and development every year for teaching and learning in science attests to the importance of reforming practice to improve student learning and participation. What many of these programs attempt to do is to build teaching expertise using new resources, new knowledge, and new skills. At the same time, it is widely known that situational issues like socioeconomic stress, lack of adequate resources, and few professional development opportunities (PD) (e.g., Ingersoll & May, 2012) make the work of teaching, especially in science, challenging to do. How can teachers successfully balance the calls for reform and the needs to consider their teaching contexts? A look into the adaptive expertise literature may provide insights into how to do this.

There is a robust tradition on the study of expertise in the learning sciences (e.g., Bereiter & Scardamalia, 1993) and teaching (e.g., Berliner, 2001). In our research, we are interested in understanding what expertise might look like and how it can be developed with new or reform-oriented curricula and instruction. How to enhance performance on novel problems has been characterized as adaptive expertise (Barnett & Koslowski, 2002; Scardamalia & Bereiter, 1993). Adaptive expertise goes beyond simply having more years of experience or possessing more knowledge in a content domain. The complex nature of teaching requires teachers to be able to orchestrate myriad of often unobservable variables, see multiple perspectives, recognize problems, and identify possibilities in existing and emergent situations (Bransford et al., 2005; Fairbanks et al., 2010). In other words, teachers need to constantly be adaptive with new or non-routine events. Adaptive experts actively seek to extend their capabilities and are always working at the edge of their competence. In this way, as Bereiter and Scardamalia (1993) suggest, expertise should be understood as a *process* rather than a *state*. Thus, searching for qualities that represent what adaptive experts look like necessarily requires us to identify actions.

With respect to actions, a review of the literature reveals three important characteristics: flexibility, ability to demonstrate deeper level understanding, and deliberate practice. Berliner (2001) writes that compared to non-experts, experts are more flexible, opportunistic in their planning, and can change enactments faster when it is appropriate. Adaptive experts flexibly and critically apply their knowledge (Ferrari, 2002) in new situations (Bransford, 1999) and are constantly learning while doing it (Chi, 2011). Flexibility is also manifested in teachers' abilities to integrate aspects of teacher knowledge in relation to the teaching act while responding to their specific contexts (Tsui, 2009). In terms of the characteristic of deeper level understanding, experts represent problems in qualitatively different ways than novices and can recognize meaningful patterns faster (Berliner, 2001). Routines are automatized, which frees up cognitive load to attend to deeper level problem solving or other tasks (Bransford, 1999; Hammerness et al., 2005) and allows experts to perform at a higher level (Ferrari, 2002; Ericsson et al., 2006). Experts are also recognized by their ability to engage in reflection and conscious deliberation (Tsui, 2009). They are highly motivated, self-regulated, and constantly seek to improve performance by identifying problems, addressing them, and finding new problems to work on (Bereiter & Scardamalia 1993; Berliner, 2001; Tsui, 2009).

In the next section, we operationalize these descriptions of adaptive expertise by applying them to three cases of teachers we worked with on a project aimed at implementing computer supported complex systems reform curricula in high school science.

Methodology

Context

Our NSF-funded project engages teachers and students in learning experiences that build knowledge of scientific practices using computational models and knowledge in complex systems and biology. The project entails building a curricular and instructional sequence in five high school biology units. Participants use biology simulations built on an agent-based modeling platform that combines graphical blocks-based programming with a 3-D game-like interface. We have developed classroom curricular materials that lay out the scope and sequence of two to three day units in each of the biology topics and constructed PD experiences during the summer and school year to support teachers' implementation of the project. All project activities are underpinned by a complex systems curriculum and instruction framework that includes: i) curricular relevance in 21st century needs, standards alignment, and collaboration between researchers and teachers; ii) cognitively-rich pedagogies instantiated in social constructivist strategies such as argumentation and constructionist learning for building computational models; iii) tools for teaching and learning including agent-based computer models and student and teacher curricular guides; and iv) content expertise in complex systems, biology, and computational thinking (see Yoon et al., 2013). Teachers are expected to integrate project units into their regular Biology curricula throughout the year. During the implementation, teachers have access to project facilitators who assist them in their classrooms; they also have access to other teachers in the project community through a shared online resource tool and periodic face-to-face Saturday PD meetings. The data reported here are from the first pilot year of the project's implementation from 2011–2012.

Case Study Methodology, Participants, and Data Sources

Since the goal of this exploratory research is to provide a model that defines and illustrates characteristics of adaptive expertise, we have chosen a case study methodology to provide rich descriptions of enactments in classrooms. Case studies enable investigators to explore multiple bounded systems over time through in-depth and multiple data collection and analyses (Creswell, 2007). Furthermore, they afford researchers perspectives on the impact of project activities in naturalistic settings (Creswell, 2007) whereby evidence can be collected and interpretations can be made based on how events unfold in real-world scenarios.

The three cases in the study have been constructed around three participants with fairly differing teacher level and school level demographics. The first participant was Imelda, who had taught for 16 years. In terms of school demographics, her school had the highest achievement of the three with 47% of students scoring in the advanced range in science and just 6% registering on free and reduced priced lunch. The second participant was Jenny, who had 9 years of high school teaching experience and advanced levels of content knowledge (the year after she participated in the project, she was accepted into the Ph.D. program at Harvard to study biology). Of the three schools, her school had the lowest achievement levels with just 14% of students in the advanced range in science on the state test and a high proportion (58%) of students on free and reduced priced lunch. The third participant was Matt, who at the time of the project had 6 years of high school teaching experience. In his school, 20% of students scored in the advanced range in science with 30% of students on free and reduced priced lunch. Across the three cases, the classes were taught at the freshman and sophomore levels.

We collected information from five data sources: i) classroom observations that averaged in total 18 hours for each teacher; ii) exit summer PD and post-implementation interviews which probed teachers' knowledge, and beliefs about the project, their contribution, and perspectives on success of the implementation; iii) teacher surveys; iv) a focus group interview with project facilitators that captured informal insights into how the implementation in each class occurred; and v) student learning outcomes as measured by growth in student understanding of complex systems in an open-ended question on ecological systems.

Data Analysis

Transcriptions of the data sources were analyzed qualitatively for instances which demonstrated adaptive expertise in the three categories of *flexibility*, ability to demonstrate *deeper level understanding*, and *deliberate practice*. The definitions of the categories were derived from the literature review of adaptive expertise. Levels of expertise were identified through an iterative mining of participant observation and interview data of the larger cohort of project participants. Examples that appeared to be upper and lower anchors of each category were discussed and agreed upon by the research team and used to construct the coding manual found in Table 1. Moderate or medium levels of each category were assigned if they fit in the category between the low and high anchors. The medium codes will eventually be added to the coding scheme with further iterations of analyses. While external validation of the coding manual is forthcoming, we attempted to achieve internal validity by assigning the analysis of each case to two members of the research team who did the coding independently and then compared results. Any discrepancies in the codes were negotiated until consensus was reached.

Table 1: Categorization Manual for Adaptive Expertise.

Category and Definition	Examples
Flexibility This category is defined by actions of the teacher that demonstrate the ability to incorporate project activities into their daily practice. Demonstrating flexibility includes an awareness of their student population and needs, an awareness of the school context, and an ability to respond to unexpected issues that arise through implementation. They are able to adapt their practice by incorporating project expectations in their situated context. (Simply eliminating parts of the curriculum is NOT flexible because there is no consideration of how to adapt within the constraints).	**High**: Teacher instructs the students that when they get to [argumentation], they need to stop and talk with another group and they should turn off their monitors when they are ready to discuss. (*Here the inference is that he understands that students will continue to play with the simulation without taking the argumentation seriously and then puts instructions into place that will force productive behaviors*). **Low**: Teacher informed us that 17 of the 30 students in this class were failing. He had also NOT prepped the students at all in terms of introducing them to the simulation environment beforehand. I don't think he reviewed the evolution concepts beforehand either. (*Despite the fact that students are failing in this class, the teacher makes no attempt to scaffold project activities for them*).
Deeper level of understanding This category is defined by actions of the teacher that demonstrate their ability to go beyond what they are required to do with the project. They have assimilated the information and are able to implement extensions or make connections that build or address deeper level of knowledge construction or problem solving. They are able to	**High**: Teacher told the students who finished early to "mess with what you've done." The students were able to create their own mini experiments by changing one of the variables in the model. This kept the group engaged while waiting for slower students. This also allowed the students space to explore and ask new questions. (*The project at this point did not require students to go under the hood and program. He recognized that for advanced students this might be an interesting thing to do*).

bring in variation from outside the present system of activity (inclusive of the classroom and school context and the project).	**Low**: From here on, the teacher is completely hands-off. It is clear some students are very confused and don't know where to start. (*He is unable to go deeper with his students*).
Deliberate practice This category is defined by actions of the teacher that demonstrate their ability to show motivation, focus, and repeated effort to monitor their practice, and to devise and subsequently attempt new approaches to improve implementation. Teachers exhibit explicit evidence of reflecting on a problem and how to improve.	**High**: Teacher was surprised in the first class that the students didn't "fly though it" and make it through the whole activity. He said that he could have prepared a bit better (*making it his responsibility to organize class better*). "You lose some of your game if you don't practice." (*Here he is metacognitive about what needs to be improved in his practice*). **Low**: In debrief with the teacher: He did not prepare his students before implementing the evolution activity. He seemed to think it 'went OK' and reiterated that this a tough group of students. (*Here he is not metacognitive about what needs to be improved in his practice and instead blames students for being "tough"*).

Case Study Results

In this section, we present instances that emerged from the study data sources that illustrate the case teacher's levels of adaptive expertise in each of the three categories. The cases are presented in a semi-hierarchical order of low adaptive expertise to medium/high. The instances selected are representative of the average level of the teacher over time although there may be other instances that diverged from those levels not reported here. In the pilot year, we did not see any teachers with solidly high levels of adaptive expertise, which we expect would improve as teachers continue to implement project activities in following years.

Imelda

Over all, Imelda demonstrated lower adaptive expertise characteristics than the other teachers. In terms of *flexibility*, she had some challenges in effectively adapting her practice and responding to students' confusion about particular aspects of the simulations. During one classroom observation, it was noted that there was "confusion regarding what the curds [were]. It was written on the board, but the students jumped in without reading it" (10/4/2011). Here, it seemed that Imelda had anticipated students' difficulties in identifying parts of the simulation by writing a key on the board beforehand to help students organize these pieces. However, the effectiveness of this scaffold was limited as she failed to direct students' attention to the key. Observations of the same unit revealed that she attempted later to alert students to the key, but that this too resulted in limited impact as "some eyes [were] still on the computer/worksheets [and therefore] not sure how many people followed" (10/5/2011). This example demonstrates Imelda's low flexibility because it illustrates her inability to adapt her instructional practices to set students up to work with and navigate through the simulation. Even when she tried to respond to students' needs, her redirection lacked the facility to effectively support students.

Imelda also demonstrated a low level of *deeper understanding* of the project's tasks and goals. This was most evident in her inability to teach parts of the project's curriculum related to her understanding of how complex systems were instantiated in the computer simulations. In classroom observations, Imelda appeared to ask a lot of simple to answer questions with no active student discussion and no student questions probing more deeply into the material. Her lack of understanding of the content prohibited her from delving into conversations with students that would help them acquire a deeper understanding of the information. For example, when students asked her about how to answer the complex systems questions in the student activity packet, observation notes state that she "deferred to [the facilitator]" (10/6/2011) to answer the questions. Despite the fact that the summer PD focused a good deal of time on definitions of complex systems, how characteristics were built into the simulations, and examining resources meant to improve teacher knowledge, evidently Imelda's inability to assimilate the information herself prevented her from extending it to her students.

In the third category of expertise, Imelda showed a moderate level of *deliberate practice*. In her post-implementation interview reflecting on how different students progressed through the activities, she explained,

> What I saw was…how different stages are for each topic we cover. There were some students who finished the whole thing and there were some who just got to the 2nd section…Given each class there were students at all different levels, putting them in smaller individual groups helped me understand where they were. But I found a lot of students moving ahead just to finish the page and move onto the next, I'm not sure if they understood what they were doing. I didn't have time to go over all the answers in class. (04/2012)

Observation notes showed that Imelda tried different configurations of student activity groups through different grouping arrangements and organization of desks. In response to seeing the disparate paces at which students completed their work, she purposefully placed her students in smaller groups to better monitor their progress. This reflects a concerted effort on her part to understand students' levels of progress, which locates her at a

moderate level of deliberate practice. However, she stopped short of improving her practice, in this example, by not following up on whether the smaller groups made a difference in her ability to monitor student learning.

Jenny

In contrast to Imelda, Jenny demonstrated higher levels of expertise in her classroom implementation. With respect to the category of *flexibility*, it was evident from observations that she needed to be flexible in being able to navigate how she would use the computers available to her in the school. One observation showed that she made room switches with her colleagues in order to use the school computer lab rather than the laptop cart, which held computers that were not compatible with the project's software. In this example, she showed even further flexibility in terms of dealing with the limitations of the computer room set up. In a study survey Jenny commented, "The challenge with the computer lab is that it is nearly impossible to hold class discussions in that room due to the layout. I overcame this by doing discussions before and after the activities, in our regular classroom, and only very infrequently pausing the students while they were working in the lab" (7/12/2012). She identified that, for her students to be able to learn from a group discussion, the layout of the room needed to be different, thus she was able to shift when and where she conducted discussion. However, while Jenny showed flexibility around adapting her practice for optimal computer use, she was only moderately flexible with the content of the curriculum. Observation notes showed that the level at which she pitched her instruction was somewhat higher than her students could comprehend. Her deep content knowledge and ability to connect content to complex systems was the highest of the three teachers and yet, often, she dwelled on aspects of the science that were not necessary for students to understand the simulation.

With respect to demonstrating *deeper level understanding*, Jenny exhibited challenges in helping her students understand how simulations and computational thinking were related to complex systems or biology. The following excerpt from observation notes on an introduction to complex systems illustrates this point.

> [Jenny] ran the simulation a few times and made various changes, explaining some parts of the code and engaging the students to consider "what if" questions. A few students actively responded but a few others started commenting out loud "what does this have to do with biology?" "What does this have to with science?" "Isn't this computer science?" (11/15/2011)

In this episode, Jenny did not appear to be able to answer student questions even though she intentionally tried to bring in computational thinking by posing "what if" questions. The fact that she did however, try to bring computational thinking into the science curriculum shows that she exhibited a moderate deeper level of understanding.

Finally, Jenny showed a high level of *deliberate practice*. In her exit interview after the summer PD, Jenny expressed concern that she would not be allowed to mold the units to fit her students' needs. She asked, "So one of my questions is to what extent, if any, am I allowed to tweak in this year? Like we have these lesson handouts that are on the flash drive that presumably are already used. I mean I always want to modify things just because of my own particular students or whatever, and I don't feel like I have a good sense at this point about that" (08/2011). In thinking about the upcoming school year, Jenny was concerned that she would not be allowed to tailor the reform activities to her students needs. It is evident that during the summer PD, Jenny was already thinking about her classroom practice, which continued into the school year. For example, after one of the classroom implementations of the enzyme activity in the Human Body unit, Jenny reflected,

> [The] Potato + hydrogen peroxide lab (which they did before this) is LESS abstract than this. But maybe this was just boring because they already knew so much (or at least had so much exposure) from that experience. For sugar transport, let's try the [project simulation] activity earlier in the sequence to see if that makes it more engaging. (11/18/2011)

In this comment, Jenny reflected on the students' experience and began already planning how to create a better learning environment for her students with the next unit. Here she was motivated and focused on improving the learning experience for her students and demonstrated a high level of deliberate practice.

Matt

Similar to Jenny, Matt exhibited moderate to high levels of expertise. In the category of *flexibility*, he showed moderate ability to adapt project materials into his daily practice. This was most prominent in his attempts at working with the students to complete the required activities in each unit despite other issues getting in the way. For example, when time constraints were an obstacle, rather than simply not completing the activity or eliminating certain portions of the lesson, Matt was open to rearranging the class schedule. One observer noted, "The lab took longer than Matt had wanted, so he changed the plan and said they should complete part 2 for today. He said he would probably need to book the lab for another period to let them complete the whole thing"

(11/15/11). Matt was not only flexible in the amount of time required for the activities he also worked to ensure every student participated. He stated in a post-implementation survey,

> This [computational thinking] is not a way in which students are used to thinking. Some students can become frustrated and confused when asked to think in a new way. As a consequence there is a balance that needs to be found to help enhance some students with computational thinking, while not losing the other students. (6/20/2012)

However, occasionally Matt did not seem to be successful at this. For example, an observer reported, "many [student] pairs having only one student working and the other not paying attention. There were also numerous distractions, with students walking through the computer lab and stopping to talk" (11/16/2011) These examples of Matt's inconsistent successes to adapt to external obstacles such as student behavior and time constraints are evidence of Matt's moderate flexibility.

In terms of the category of *deeper level understanding,* Matt showed moderate ability. Several observations noted that he contextualized the project lessons by embellishing the content with demonstrations, "Matt did two demonstrations about diffusion: match (odor) and dye-in-water (liquids)," and questioning, "Matt pointing graphically to the different simulation shapes and asking questions about what components are" (11/10/2011). He also led the class in discussions before and after the lesson each day. For example in one lesson, "Matt did a wrap-up as a class: He summarized: 'the first run was random, was any one color always dominant?" This session continued with a class discussion. Matt occasionally circulated the room, answering students' questions, probing student understanding, and further explaining concepts (4/11/2011), but as one facilitator noted in the focus group interview, "Just comparing him [Matt] to the other teachers, he was much more hands off. He'd hand out the worksheets. He'd actually contextualize the activity, which is nice because not all the teachers did that but then he just kind of let the students go" (5/18/13). Matt's incorporation of simple demonstrations, asking and responding to questions, and leading discussions that supported lesson content are evidence of deeper understanding of the goals of the reform, but because these efforts lacked more robust biology and complex systems content and his hands-off approach during the activities indicate that he did not necessarily facilitate student learning on a deeper level beyond basic conceptual understanding.

Regarding *deliberate practice*, facilitators viewed Matt as one of the most reflective teachers in this cohort. As one facilitator reported in the focus group interview, "He was the most open to feedback of any of the teachers and spent the most time with us debriefing during and after implementations on how he could improve and how he could change things and how we could change things to make them better for him" (5/18/13). His intention to continue honing his approach to incorporating the reform curriculum was clear. On the post-implementation survey at the end of the year he wrote,

> In the future I plan on being more conscious of complex systems in biology and expressing and exploring these ideas more explicitly in my teaching…I hoped that themes of randomness, complex systems, and emergence would carry over into other parts of the curriculum that did not happen to the degree that I hoped, but still intend to pursue this further, next year. (6/20/2012)

While there is no evidence of the frequency and consistency with which Matt practiced deliberately, such statements indicate his interest and intent to improve his practice, which were higher than the other two case teachers.

In the next section, we provide a summary of our efforts in this research. We discuss implications for using an adaptive expertise model as a lens to understand classroom enactments of computer supported complex systems reform curricula with the goal of improving teachers' instructional practices.

Discussion

We were motivated to do this research for several reasons. The study of complex systems has become a prominent research focus in the science and social sciences (The National Academies, 2009) with equally increasing emphasis on learning and instruction particularly in the recently constructed *Next Generation Science Standards*. Although curricular and computational resources to support the learning of complex systems with students has been a focus in the learning sciences, relatively little attention has been paid to understanding how teachers can best support classroom learning experiences, how they can become experts in this support, and what this expertise might look like. Acknowledging the importance of addressing the situated nature of teachers, classrooms, and schools (Penuel et al., 2011), in this paper, we have proposed an adaptive expertise model for evaluating implementation experiences that aims to illustrate the differential qualities of teachers' abilities to adapt complex systems curricula and tools in their biology courses. The model outlines three important characteristics of adaptive expertise: *flexibility* in accommodating project activities into daily classroom and

course requirements, ability to demonstrate *deeper level understanding*, and intending or showing efforts toward *deliberate practice* to improve instruction.

The examples above do not provide evidence of every level of expertise within each category defined by the model, but there was sufficient data to rank the teachers' levels of adaptive expertise relative to one another. Ultimately this overarching assessment was determined not only by quality of expertise exhibited by each teacher, but by the frequency and consistency of such observations. For example, Imelda showed little to no effort to incorporate project activities into her classroom. Jenny exhibited high levels of flexibility at times, but only with respect to technology use, and this flexibility was not consistent. Although there are several instances in which Matt's flexibility was high in his response to student engagement or time constraints, he demonstrated some hands-off behaviors at times when students may have needed some further scaffolding. Regarding exhibiting deeper levels of understanding, Imelda's implementation of the lessons was limited and rarely deviated from the teacher's guide. Although Jenny's biology content knowledge was extensive, she often over-reached in terms of students' abilities to comprehend the science. Matt most frequently contextualized the lessons to make the science accessible to students though his efforts were somewhat superficial. All three participants shared reflections regarding their teaching practice, but Matt also initiated dialogue with and sought feedback from project facilitators in order to further reflect upon and improve his practice.

We believe that using an adaptive expertise model helps professional developers and researchers interested in learning how to train teachers to teach with complex systems resources and approaches (Hmelo-Silver & Azevedo, 2006) by illustrating the range of contextualized classroom enactments. As Fairbanks et al. (2010) and others suggest, the complex nature of teaching requires the orchestration of many variables and responses to emergent situations that often cannot be predicted a priori. Particularly with computer-supported learning, where teachers' knowledge of and exposure to technology can impact integration success, (Aldunate & Nussbaum, 2013; Ertmer et al., 2012; Mueller et al., 2008), understanding why or how integration is challenging can be instructive in PD activities. Furthermore, by investigating and identifying instances in practice that illustrate the characteristics of an adaptive expertise continuum, we aim to articulate a *process* rather than a *state* (Bereiter & Scardamalia, 1993) that can actively extend teachers' capabilities beyond their current competencies. From our small sample of cases, we have already begun to construct tools for developing teachers' adaptive expertise. For example, we have matched teachers who have the same school demographics and provided release or substitute teacher funds for teachers to conduct peer observations. We have selected teachers who have exhibited higher levels of expertise and arranged videotaping of a project unit in action for other teachers to view and discuss. During our most recent summer PD workshop, we invited teachers to lead discussions about challenges in implementation, and we structured multiple peer sharing sessions for teachers to be able to listen and learn about alternative instructional approaches. These sessions were even extended to peer construction of differentiated lessons for English Language Learners and students with special learning needs. This example specifically encourages deeper level problem solving (Bransford, 1999; Hammerness et al., 2005) that ideally will enable our teachers to perform at a higher level (Ferrari, 2002; Ericsson et al., 2006).

Another area of interest for this research involves providing authentic examples of teachers in action who are working with these reform oriented complex systems curricula. Although we aim to articulate a model of adaptive expertise, we also aim to use the model to map teacher practices to provide other researchers with modular images and descriptions that may be expected of their participants in such projects. Knowing the range can help researchers intentionally plan PD activities that will best support all teacher learners from the start. We are also ultimately interested in seeing whether our categorization of adaptive expertise correlates with, or even better, predicts student-learning outcomes. In this area, we have already begun to compare the three cases of teacher expertise to the amount of student gain of complex systems understanding which was measured through open ended short answer responses to questions on ecological systems. We hypothesized that findings would reveal a link between higher levels of adaptive expertise as defined by this framework and higher student learning outcomes. Interestingly, across the three teachers, Imelda's students showed negative significant growth in complex systems understanding ($F(1,16) = 6.572$, $p = .012$). Jenny's students showed no growth ($F(1, 6) = 3.769$, $p = .110$) and Matt's students showed marginal significant growth ($F(1, 53) = 3.656$, $p = .061$). We are aware of the unevenness of sample sizes in each of the classrooms, which is why we have not offered these results as official evidence at this point. But the student learning outcomes analyses serve as some encouragement that our adaptive expertise model may be able to help in assessing the success of teachers' contextualized implementation of complex systems resources and approaches. Currently, we are analyzing more data with a larger number of teachers and students to see whether this hypothesis is true.

References

Aldunate, R. & Nussbaum, M. (2013). Teacher adoption of technology. *Computers in Human Behavior*, 29, 519-524.

Barnett, S., & Koslowski, B. (2002). Adaptive expertise: Effects of type of experience and the level of theoretical understanding it generates. *Thinking and Reasoning*, *8*(4), 237-267.

Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise.* Illinois: Open Court.

Berliner, D. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35,* 463-482.

Bransford, J., Darling-Hammond, L., & LePage, P. (2005). Introduction. In L. Darling-Hammond & J. Bransford (Eds.), Preparing teachers for a changing world: What teachers should learn and be able to do (pp. 1-39). San Francisco: Jossey-Bass.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). How experts differ from novices. In Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). *How People Learn: Brain, Mind, Experience, and School* (pp. 19-38). Washington DC: National Academy Press

Chi, T. (2011). Theoretical perspectives, methodological approaches, and trends in the study of expertise. In Y. Li and G. Kaiser (Eds.), *Expertise in mathematics instruction: An international perspective* (pp. 19-39). New York: Springer.

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* Sage Publications, Inc.

Darling-Hammond, L., & Youngs, P. (2002). Defining Highly Qualified Teachers: What does Scientifically-Based Research actually tell us? *Educational Researcher, 31(9),* 13-25.

Ericsson, KA., Charness, N., Feltovich, P.J., Hoffman, R.R. (2006). *The Cambridge Handbook of Expertise and Expert Performance*. New York: Cambridge University Press.

Ertmer, P.A., Ottenbreit-Leftwich, A.T., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers and Education*, 59, 423-435.

Fairbanks, C.M., Duffy, G.G., Faircloth, B.S., He, Y., Levin, B. et al. (2010). Beyond knowledge: Exploring why some teachers are more thoughtfully adaptive than others. *Journal of Teacher Education*, 61(1-2), 161-171.

Ferrari, M. (2002). *The Pursuit of Excellence Through Education.* Mahwah, NJ: Lawrence Erlbaum Associates.

Grotzer, T.A. (2012). *Learning causality in a complex world: Understandings of consequence.* Lanham, MD: Rowman Littlefield.

Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). In L. Darling-Hammond and J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358-389). San Francisco, CA: Jossey-Bass.

Mueller, J., Wood, E., Willoughby, T., Ross, C., & Specht, J. (2008). Identifying discriminating variables between teachers who fully integrate computers and teachers with limited integration. *Computers & Education*, 51, 1523-1537.

National Research Council. (2011). *Successful K–12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: The National Academies Press.

OECD (2011), Lessons from PISA for the United States, strong performers and successful reformers in education. OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264096660-en

Pierson, M.E. (2001). Technology integration practice as a function of pedagogical expertise. *Journal of Research on Technology in Education*, 33(4), 413-430.

Salas, E., & Rosen, M. (2010). Experts at work: Principles for developing expertise in organization. In S. Kozlowski and E. Salas (Eds.), *Learning, training, and development in organizations,* (pp. 99-134). NY: Routledge.

Tsui A.B.M. (2009). Distinctive qualities of expert teachers. *Teachers and Teaching: theory and practice*, 15(4), 421-439.

Yoon, S. A. (2011). Using social network graphs as visualization tools to influence peer selection decision-making strategies to access information about complex socioscientific issues. *Journal of the Learning Sciences*, *20*(4), 549-588.

Yoon, S. (2008). An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education*, *30*(1), 1–32.

Yoon, S., Klopfer, E., Wang, J., Sheldon, J., Wendel, D., Schoenfeld, I., Scheintaub, H., & Reider, D. (2013). Designing to improve biology understanding through complex systems in high school classrooms: No simple matter! *In the proceedings of the Computer Supported Collaborative Learning*, Madison, Wisconsin.

Acknowledgments

This research is supported by a grant from the U.S. National Science Foundation under the program *Discovery Research K-12* (DRL 1019228). Many thanks to our other project collaborators, Josh Sheldon (MIT), Ilana Schoenfeld (MIT), Daniel Wendel (MIT), Hal Scheintaub (Governor's Academy), and David Reider (Education Design, INC).

Promoting 5th Graders' Views of Science and Scientific Inquiry in an Epistemic-Enriched Knowledge-Building Environment

Feng Lin, Carol K.K. Chan, Jan van Aalst, The University of Hong Kong, Pokfulam, Hong Kong
Email: irisfeng83@gmail.com; ckkchan@hku.hk; vanaalst@hku.hk

Abstract: This paper reports on an ongoing study that examined the design of a knowledge-building environment that integrates knowledge building and epistemic change theories in fostering conceptual and epistemic growth. The study uses a quasi-experimental design in which four classes of 5th graders in Hong Kong participated in a unit focusing on electricity. The intervention involved students' collective inquiry and, epistemic reflection on what science is about; students used Knowledge Forum® to collaboratively work on pursuing ideas; and students reflected on scientific progress with their own knowledge building inquiry. A scheme was developed in assessing students' views of science; and results showed significantly stronger effects for the knowledge-building group compared with the regular inquiry group on epistemic and conceptual learning. Regression analysis showed that students' forum engagement and epistemic views contributed to their posttest conceptual understanding over and above prior science knowledge. Qualitative analysis suggested how students' experience in knowledge building might shape students' understanding about the nature of science and improve their conceptual understanding.

Epistemic cognition, or thinking about the nature of knowledge and knowing, has received much research attention in recent decades. Substantial evidence has shown that sophisticated personal epistemologies are important predictors of learning, including learning strategies and processes and conceptual change (Hofer & Pintrinch, 2002). Of particular interest about students' views of knowledge and knowing relates to what they think about the nature of science and how scientific knowledge is created. Although there is widespread recognition of the importance of scientific inquiry, often students do not think of science as an epistemic idea-driven and theory-building process. Rather children tend to think of science as concrete activities, and scientific inquiry as acquiring sets of skills such as the methodical collection of data and testing of variables (Chinn & Malhotra, 2002). The purpose of this study is to investigate students' epistemic understanding of science as a theory-building process in the context of their working on knowledge building inquiry supported by Knowledge Forum®, a computer-supported collaborative learning environment. Specifically, the goal is to design and examine a learning environment to foster students' epistemic and conceptual growth via linking views of science with knowledge building inquiry, and also to investigate the intertwined relationship between the designed environment, epistemic cognitions and conceptual understanding.

Theoretical Perspectives

Epistemic Cognition and Views of Science

There are different research traditions that examine what people think about the nature of knowledge and nature of science. Some extended the psychometric tradition initiated by Schommer (1990) and Hofer & Pintrich (1997) and took a multidimensional approach to examine epistemic beliefs; examples of the examined dimensions included certainty of knowledge, source of knowledge, justification for knowing, and development of knowledge (Conley et al., 2004). Another major approach developed from research in science education. For example, Lederman and colleagues (2002) examined several aspects of the nature of science (NOS), such as the empirical nature of science, tentative nature of science, creative and imaginative nature of science, and inferential nature of science, and so on. A third tradition examines students' epistemology of science from the "role of idea" perspective. Herein, science is perceived as a theory building process and the construction of ever-deeper explanations of the natural world (Carey et al., 1989; Chuy, et al., 2010). Carey et al. (1989) developed a clinical interview protocol and identified three general levels of understanding about science among middle school students, ranging from seeing science as discovering facts and making inventions, to seeing it as constructing explanations for phenomena. Later Smith et al. (2000) extended this line of research and elaborated science as theory building, and Chuy et al. (2010) further developed the interview protocol to examine children's understanding of science in relation to knowledge building. This study follows this tradition (Bereiter et al., 1996) and examines students' understanding of science as an idea-driven and theory-building process. We choose to build on this line of research mainly for two reasons: (1) science as a theory-building process (Kuhn, 1970) is important but often overlooked and portrayed merely a process of observation and experimentation in school science (Carey & Smith, 1993); (2) how students understand the theoretical progress of science may be facilitated as they build knowledge together; and such research direction may help develop new instruments and

extend our understanding of epistemology in relation to community process. We use the term epistemic cognition to refer to a broad notion that encompasses both cognitions of the nature of knowledge (Hofer & Pintrich, 2002) and nature of science (Lederman et al., 2002).

Research has shown that epistemic cognition influences students' conceptual change (Mason, 2000; Qian & Alvermann, 1995), for example, epistemic cognition may affect learners' intention to restructure knowledge (Vosniadou & Brewer, 1987). However, many previous studies conceptualize the relationship between epistemic cognitions, thinking and learning within theories of knowledge construction within individual learners. As learning theories have shifted from individual to collective, knowledge construction is no longer perceived as an individual but a social process (Brown et al., 1989). There is a need to extend investigation of epistemic cognition of scientific process to socio-cognitive processes and to examine how it may influence cognition in social context. As well, scientific progress and theory building evolves in scientific communities via collective advances, not only individual endeavors. Thus far, there are few investigations looking into the social and collaborative aspects of epistemic cognition. This study will examine students' views of science as a social-cognitive and community process, and investigate how these views relate to students' knowledge building and conceptual understanding.

Computer Supported Knowledge Building, Reflection, and Epistemic Change

Knowledge building is an educational model that has attracted much research attention in learning sciences and CSCL (Scardamalia & Bereiter, 2006). Its role in student learning and cognition has been discussed in a growing number of studies (e.g., Lee et al., 2006; Zhang et al., 2007; van Aalst & Chan, 2007). Knowledge building emphasizes on students, similar to scientists, working as a community and taking collective cognitive responsibility for idea improvement (Scardamalia, 2002). In knowledge building, students' collaborative discourse is supported by a computer supported collaborative learning platform, Knowledge Forum®(KF), in which students pose questions, make conjectures, co-construct explanations, reorganize ideas, and revise and integrate ideas. To make knowledge building more explicit, knowledge building principles (Scardamalia, 2002) have been proposed to guide students and teachers' knowledge building practice. These principles provide epistemological scaffolds for students' knowledge work in their community (usually their class), as students work with the principles, that might help them move from naive epistemology to more sophisticated one. Primarily, "Knowledge building is not just a pedagogical approach but a theory of epistemology" (Bereiter, 2002). The knowledge building principles, e.g., improvable ideas; rise above; constructive use of authoritative sources, and their technological function on Knowledge Forum, that is, making ideas explicit and as subject for building on and revision, have important epistemic indications for students.

The influence of knowledge building and metacognitive reflection on students' conceptual understanding has been examined (van Aalst & Chan, 2007), and the important relations have been identified among epistemic cognition, knowledge building, and conceptual understanding (Chan & Lam, 2010). Research on science and epistemic inquiry has examined the roles of scaffolding epistemic standards and scientific practice (Sandoval et al., 2005), however, fewer studies have attempted to design epistemic-enriched knowledge-building environments helping students to reflect on their knowledge building practice that mirrors the theory-building nature of science. Research has indicated (Carey et al., 1989) that if students are to understand the role of theory in science, they need to be engaged actively in the explanation-based theory-building process and make metacognitive reflections about the process. As knowledge-building model emphasizes idea improvement and collective advances, it provides a rich environment to understand and to foster students' understanding of the nature of science from the perspective of theory building. Accordingly, we aimed to design a knowledge building environment, emphasizing the theme of students working as communities and focusing on epistemic standards and the practice of scientists. A key idea is that when students working on ideas in knowledge building, they are better able to experience the role of idea and theory building in science, which would in turn foster their epistemic views and conceptual understanding.

To iterate, this study aimed to design an epistemic-enriched knowledge-building environment to foster more sophisticated epistemic cognitions and conceptual understanding among students. Three research questions were addressed: (1) Do students engaging in epistemic enriched knowledge building achieve more epistemic and conceptual growth than students in a regular inquiry-based learning environment? (2) What are the relationships among knowledge building, reflection, epistemic cognition, and conceptual understanding? (3) How do the designed environment help students improve their epistemic cognition and conceptual understanding? This paper reports preliminary findings for these questions.

Methods

Participants and Procedure

102 5th graders (age ranging 10-11) in four science classes in Hong Kong participated in this study. The experimental group included two classes, Class 1 (n=33) and Class 2 (n=19), engaging in computer-supported

knowledge-building inquiry with epistemic reflection. The comparison group also included two classes, Class 3 (n=26) and Class 4 (n=24), taught with regular inquiry-based approach with metacognitive reflection. The experimental classes were selected according to teachers' knowledge building experience, as two teachers (Class 1 & Class 2) have used knowledge building pedagogy for 4 years, and the other two (Class 3 & Class 4) are familiar with general inquiry-based approach. The study was conducted in six sessions lasting for about three weeks. All four classes worked on an extended topic from their textbook study of electricity. These four classes have similar access to learning resources (video, experiment equipment, reading material) except that the experimental classes used knowledge-building design that involved Knowledge Forum and epistemic reflection.

Designing a Knowledge Building Environment for Epistemic and Conceptual Change

The experimental environment was designed based on an integration of knowledge building pedagogy (Chan, 2011) and epistemic change theory (Bendixen, 2002). Primarily we engage students in a knowledge-building environment to pose problems, questions and explanations, and to make their ideas improvable, and to advance community knowledge as in scientific communities. We also adapted epistemic change model that emphasizes epistemic doubt and resolution of doubts, enriched in a community of knowledge builders. Students work on collective inquiry as communities of little scientists. The specific design principles are described below:

(1) *Articulate and activate prior understanding*. Before engaging in inquiry, students were asked to write down what science is about as well as their understanding about electricity, so that their ideas can be made visible in the class and open for revision. This was designed not only for creating a knowledge building culture where ideas can be examined, but also for promoting students' epistemic awareness and triggering their epistemic doubt by learning about their peer's different epistemic theories.

(2) *Start inquiry with authentic problems*. We first provided students with everyday situation (video on lemon juice and salt water conducting electricity), to stimulate their wonderment about conductors. Then they wrote out their questions and ideas on Knowledge Forum based on this material. Students also worked together in groups to test out the conductivity of different materials. Inquiry-based activity and Knowledge Forum were intertwined: students continually worked on Knowledge Forum after the experiment. Scaffolds were provided on Knowledge Forum to help them build and revise theories and explanations: "I need to understand", "my theory (explanation)", "new information", "a better theory (explanation)", "your theory cannot explain".

(3) *Deepen inquiry through knowledge building talk and experiment.* In order to facilitate good knowledge building discourse and engage them in deep construction of knowledge, knowledge-building principles (e.g., improvable ideas, epistemic agency) were explicitly discussed in the class, and linked to scientific process and scientific community. To test the ideas discussed on Knowledge Forum, students worked in groups to design experiments and make posters. Scaffolds were provided to facilitate this collective inquiry process: "our question", "our theory", "our hypothesis", and "our experiment design". After testing their ideas, students continued to write on Knowledge Forum to revise their theories. Meta-views were also created to encourage students to rise-above their existing theories.

(4) *Trigger and resolve epistemic doubt through collective epistemic talk.* Classroom discourse was conducted to scaffold students toward viewing scientific inquiry as a theory-building and theory-revision process. The teachers triggered conceptual and epistemic doubts and engaged students to reflect on the experiments and evidence; students were encouraged to think like scientists as they pursued inquiry and considered the need for revising their hypotheses and theory. Online and offline discourse was linked: Students discussed 'electricity' in one view alongside another 'view' that asked them to reflect on how scientists construct knowledge, and how their work might be similar to scientists.

(5) *Understanding science as theory-building through collective epistemic reflection.* Knowledge building emphasizes collective inquiry and aims to help students to understand the social and collective nature of science. Students therefore were asked to reflect on their own scientific and knowledge building inquiry process as "little scientists." Experts' inquiry process and epistemic theories were illustrated on a worksheet pertaining to four different models of scientists, and students were asked to reflect and identify the part that they think they have experienced. This was a scaffold designed to help students connect what they do and what scientists do. Teachers then promoted a discussion among students to investigate the similarities between their own collective inquiry process and the social construction process in scientific community.

The comparison classes went through similar processes as the experimental classes. They first wrote their prior conceptions about electricity and nature of science, and watched the same video to start inquiry. They did the same experiments on conductors, and worked in groups to design experiments to test their own ideas. The scaffolds provided in the group inquiry were same as the experimental group. The only major difference is that the experimental group used Knowledge Forum for inquiry, and had epistemic talk and reflection. To make the comparison more equivalent, the comparison group students were asked to write reflection journals on papers after class, and were provided with metacognitive scaffold, such as "my new learning".

Measures

Written Questions on Epistemic Views of Science

All children were administered the written questionnaires at the beginning and end of the unit. Children's epistemic views of science were measured with 8 written questions developed in this study based on Carey et al. (1989), Smith et al. (2000), Chuy et al. (2010), and Lederman & Ko's (2004) items. Premised on the framework of science as theory building, four components were identified: (1) Role of idea (2) Theory building and theory revision (3) Theory-fact understanding; and (4) Social process of scientific progress. Examples of the questions included: "What is science"; "What do scientists do"; "Why do scientists do experiments"; "Scientists may have different even contradictory ideas, do you think it is good for science, and why".

Interviews on Epistemic Views of Science

8 students from each class (i.e., n = 32) were interviewed before and after the intervention with the epistemic cognitions interview protocol to examine their epistemic and conceptual change process. The first part of the interview questions were similar to the questionnaire items but allowed students to elaborate on their thinking about science; the second part had questions that mapped with the first parts but probed students' understanding of their own inquiry process. The third part of the interview asked students to reflect how they had changed their views of science and conceptual understanding about electricity. These data are currently being analyzed.

Conceptual Understanding

Students' conceptual understanding was measured with a knowledge test containing different parts: The first tested students' understanding about the conductivity of different materials (e.g., metal, distilled water, juice, graphite); the second asked them to give explanation on why some materials conduct electricity, the third part asked what they knew about electricity and what questions they had. Students' responses to the first part of the test were scored according to the scientific correctness of answers, and their responses to the second and third parts were coded on a 4-point scale based on depth of explanation.

Knowledge Building Engagement on Forum

Students' knowledge building engagement was assessed with a software Analytic Toolkit (ATK) developed by the Knowledge Building Research Team at the University of Toronto (Burtis, 1998). We selected two indices to illustrate students' collaboration on Knowledge Forum: percentage of notes *linked* and percentage of notes *read*.

Epistemic Reflection on Science (Students as Little Scientists Worksheet)

As discussed earlier, the worksheet described the models and practice of four different scientists involving concrete day-to-day work as well as theory-building process of scientists. Students were asked to identify aspects that were similar to what they did in knowledge building inquiry. This worksheet provides a scaffold for students to engage in classroom discourse; and at the same time, it provided data on students' reflection of their inquiry and epistemic process. A 5-point scale was developed to code the worksheet ranging from naïve, concrete, elaborated, theory-change and social-community processes of scientific progress.

Analysis and Results

Characterization and Change of Epistemic Views of Science

Students' responses to the epistemic questionnaire were coded to identify students' epistemic cognition ranging from viewing science as making concrete materials to an idea-driven and theory-building process. There are many systems and this coding scheme is based on the theoretical framework, and in line with cognitive studies, different levels are included. Based on top-down and bottom-up analysis, a 4-point coding scheme was developed, for example, for one of the items on role of idea (what is science?), at level 1, students responses showed a rudimentary understanding focusing on concrete activities (e.g., "*Science is about inventing new things, for the convenience of people")*; at level 2, students showed some awareness of the existence of abstract unseen entity in science (e.g., "*Science is about investigating some questions");* at level 3, responses reflect some understanding about the relationship between theory and experiment, and the explanatory nature of science (e.g., "*Science is an investigation, it involves experiments; to explain all kinds of phenomenon"*); at level 4, responses indicate a deeper understanding of theory building (e.g., "*Science is about making theories through experiment, then they do different experiment to revise the theory"*). For the item on scientific progress as a social process, responses at level 1 do not appreciate the role of different ideas for scientific progress (e.g., "*it is not good to have different ideas. When you have too many ideas, it is hard to find answer.*"); responses at level 2 show some superficial understanding (e.g., *"it is a good thing, you can compare and see which one is right*); responses at level 3 appreciate of the role of idea interaction in science (e.g., "*it is a good thing, different ideas can inspire scientists*"); responses at level 4 indicated better understanding about the role of different ideas

for theory improvement/knowledge creation (e.g., "*scientists have different ideas, they test them with experiments; they may understand some new ideas...organize them.. and make a new theory"*).

To examine the intervention effects on the change of epistemic cognition over time, a 2 x 2 (group x time) repeated measures MANOVA was performed for the scores of four scales (Table 1). The participants were nested within classes and therefore the measurements were not statistically independent, which may affect Type I error rates. All alphas were therefore set at .01. Results revealed statistically significant multivariate effects; Follow up univariate ANOVAs revealed significant main effects for time for: role of idea, $F(1,100)=85.06$, $p<.001$, Partial $eta^2=.46$; theory building, $F(1,100)=37.07$, $p<.001$, Partial $eta^2=.27$; social process, $F(1,100)=8.95$, $p<.01$, Partial $eta^2=.08$; and theory-fact understanding, $F(1,100)=5.59$, $p<.05$. Partial $eta^2=.05$, suggesting both knowledge-building and inquiry groups improved over time. There was also a significant main effect for groups for role of idea, theory building, and social process. Importantly, significant time x group interaction effects were obtained on role of idea, $F(1,100)=16.89$, $p<.001$, Partial $eta^2=.14$, theory building, $F(1,100)=9.79$, $p<.01$, Partial $eta^2=.09$, and social process $F(1,100)=17.07$, $p<.001$, Partial $eta^2=.15$. The interaction effect was not significant for theory-fact understanding. Results indicated that knowledge-building group had more gains on epistemic views of science than did the comparison group.

Table 1: Pre and posttest epistemic cognitions mean scores (SD in parentheses) across classes

	Knowledge building group (n=52)		Comparison group (n=50)	
	Pre-test	Post-test	Pre-test	Post-test
Role of Idea	1.26(.33)	2.03(.66)	1.12(.32)	1.42(.49)
Theory-Fact	1.44 (.67)	1.79 (.94)	1.42 (.84)	1.54 (.73)
Theory Building	1.63(.51)	2.28(.70)	1.36(.56)	1.57(.46)
Social aspect	2.37(.95)	3.12(.94)	1.98(.87)	1.86(.83)
Epistemic overall	6.69 (1.43)	9.21 (2.07)	5.88 (1.63)	6.39 (1.71)

Changes in Conceptual Understanding

The mean and SD of the pre and post scores were 1.11 (.23) and 1.55 (.18) for knowledge building group (n=52), and .94 (.14) and 1.12 (.15) for comparison group (n=50) at pre and posttest respectively. Repeated measure ANOVA was conducted to test the intervention effect for conceptual understanding. The results showed a time effect, $F(1,99)=217.18$, $p<.001$, Partial $eta^2=.687$, and a group effect, $F(1, 99)=1.6.31$, $p<.001$, Partial $eta^2=.518$. There was also a significant time and group interaction effect, $F(1,99)=39.00$, $p<.001$, Partial $eta^2=.283$. These results indicate that both groups improved their conceptual understanding over time, but the knowledge building group had a larger gains than had the comparison group.

Relations between Knowledge Building, Epistemic Views &Conceptual Understanding

The second question investigated the relationship among knowledge building, epistemic cognition, and conceptual understanding, and examined the prediction of different variables on posttest conceptual understanding. Analyses were conducted within the knowledge building groups (n=52) as ATK indices were available only for this group. Correlations analyses indicated that students' post-test conceptual understanding was related to post-test epistemic cognition and KF-link indices. As well, students' post epistemic cognition was related to their epistemic reflection. Primarily, students' engagement in Knowledge Forum and their epistemic cognition are related to their conceptual understanding after instruction.

Table 2: Correlation among post epistemic cognition, post conceptual understanding, KF collaboration (link & read), and epistemic reflection (n=52)

	Epistemic cognition	Conceptual understanding	KF link	KF read
Conceptual understanding	.310*	1		
KF link	0.136	.304*	1	
KF read	.242	0.17	.287*	1
Epistemic reflection	.377**	0.143	.304*	0.155

Note: *p<.05; **p<.01;

Prediction of Prior Knowledge, Epistemic Cognition and Knowledge Forum Activities on Conceptual Understanding

Hierarchical regression analyses were conducted for the knowledge-building group (n=52); first entering pre-test scores, followed by Knowledge Forum activities, and then epistemic cognition. Results showed that prior conceptual understanding explains 13% of variance (R^2=.13), when Knowledge Forum note links was added, an additional 6.2% variance was explained. When post epistemic cognition was added, another further

8.3% variance was explained. These results indicated that over and above prior knowledge, students' collaboration on Knowledge Forum, epistemic cognition contributed to the post conceptual understanding.

Table 3. Hierarchical regression on post conceptual understanding with prior conceptual understanding, KF note link, and post epistemic cognition (n=52)

	R	R^2	R^2 Change	F Change	Sig.
Prior conceptual understanding	.36	.13	.132	7.59**	.008
KF link	.44	.19	.062	3.75(*)	.058
Post epistemic cognition	.526	.28	.083	5.52*	.023

Note: *p<.05; **p<.01

Prediction of Learning Context, Prior Knowledge, Epistemic Cognition on Conceptual Understanding

Hierarchical regression was also conducted to examine the contribution of posttest epistemic cognition score and learning context (group) to students' posttest conceptual understanding (n=101). The learning context was coded into two variables (KB group=1, None KB group=0). Results showed that prior conceptual understanding explained 26.7% of the variance (R^2=.267). When post epistemic cognition was added, additional 27.7% of the variance was explained. When learning context was added, additional 16.5 % of the variance was explained. Results indicated that over and above prior knowledge, epistemic cognition and knowledge building environment contributed to post-conceptual understanding.

Table 4. Hierarchical regression on conceptual change score with prior conceptual understanding, post epistemic cognition, and learning context (n = 101)

	R	R^2	R^2 Change	F Change
Prior conceptual understanding	.52	.267	.267	36.098***
Post epistemic cognition	.74	.544	.277	59.623***
Learning context	.84	.709	.165	55.122***

Note: ***p<.001

The Epistemic and Conceptual Change Process: Preliminary Observation

The third research question examined how knowledge-building environment might support the observed epistemic and conceptual change. Consistent with the quantitative findings, qualitative analysis of students' post interview data also revealed the possible positive impact of knowledge building on students' epistemic and conceptual change. Table 5 shows a comparison of two students' reflections on their own inquiry process, one from the knowledge-building group and one from the comparison group.

Table 5: An example of students' interview responses about their own inquiry process

Student A (from knowledge building group)	**Student B (from comparison group)**
I: What makes a good inquiry-based discussion?	**I: What makes a good inquiry-based discussion?**
A: Classmates will bring out questions, there will be hypothesis, theories….new knowledge is produced in the theories. In [working on] new knowledge, there are things [we] don't understand, so we will keep asking questions…. Then there are more questions, hypothesis, and knowledge, etc. It keeps circulating…	B: everyone discusses with each other.. I: any others? B: no …..
I: Is it good to have different ideas in your discussion?	I: Is it good to have different ideas in your discussion?
A: **It is the same like what I said about scientists**. It is good, with different theories, we can have better theory.	B: If you can do it separately (experiment), it is ok to have different ideas….but we were working as a group, therefore it is not good to have different ideas.
I: how can different theories help you find better theory?	
A: the different ideas will help you revise your theory, which means, your theory will become better...	I: How was new knowledge created in your class?
I: How was new knowledge created in your class?	
A: We kept discussing, brought out questions, hypothesis, theories...there may be new knowledge there..…and as I said, it is a cyclic process	B: We need to try another activity and find a correct answer, then we get new knowledge.

As suggested in these excerpts, student A (knowledge building group) seemed to have a better sense that scientific knowledge is socially constructed, and that it improves in a cyclic manner through constantly questioning and theorizing. He also appreciated the role of different ideas in improving knowledge. However, student B thought of science as activities and getting something right for new knowledge. Student A was

alluding to his experience as they worked on scientific inquiry and discussion. These excerpts provided some glimpses into how the knowledge-building environment enriched with epistemic reflection might have influenced students' views of science. Other examples are also included to show correspondence between students' epistemic understanding of science and their experience with Knowledge Forum. When asked "*how do you think scientists create new knowledge",* one knowledge building student said, *"scientists will construct a new theory based on previous theories. Just like the working on Knowledge Forum, you will write and type something under a question, and it will be revised gradually, or a new theory will be proposed.*" From this excerpt, we can see how students' knowledge building engagement and the visualization of ideas on Knowledge Forum may have helped them understand the theory construction process of scientists.

To further understand how knowledge building may have possibly provided students with an epistemic environment, we examined student's understanding of scientists' inquiry and also reflection of their own inquiry process. For example, a knowledge-building student LHT reflected: "*When working on Knowledge Forum, underneath different themes people posted questions…it was all about electricity, but we talked about different aspects of it, e.g., some were wondering why salt water can conduct electricity, some others were wondering why fruit can conduct electricity…. Underneath all these questions, there were different [ideas], some of which were theories. You might question these theories, so you asked follow-up questions. However, people might even question your follow-up questions, so you had to revise it…. Some used the scaffold "your theory cannot explain" in their responses, which means your theory might have problems. It keeps going on like this, and it becomes a big cluster of notes full of questions, theories, and questioning etc……*". While talking about her view of scientists' inquiry, this student mentioned a similar progressive process: "*….after scientists make a theory, there may be follow-up questions…." "…Theory will change with time, when other scientists make another investigation of the theory, they may find some problems in it. Then they will do experiment to test…. and it will be changed*". These excerpts provided glimpses into how knowledge-building inquiry, supported by Knowledge Forum, may have shaped students' understanding about the progressive and collective nature of science, and helped them revise their theories and explanations through its social mechanism.

Conclusion, Implications and Significance

This study designed a knowledge-building environment that attempts to integrate knowledge building and epistemic change theories, and examined the role of such design in facilitating epistemic and conceptual change in fifth graders. In the process, we developed a coding scheme to examine children's epistemic understanding of science focusing on role of idea, theory-fact understanding, theory building and social process of scientific progress. Consistent with the research of Carey et al. (1989) and Smith et al. (2000), results showed variation among these participants ranging from seeing scientific inquiry as concrete activity to viewing it as an idea-driven theory-building process. As well, these children demonstrated understanding of scientific progress as propelled by inquiry and social processes in a community. Our findings showed that grade 5 students working in the knowledge-building environment obtained significantly more changes towards more sophisticated views of scientific inquiry, compared to students in the regular inquiry-based classes; and they also obtained higher scores on conceptual understanding of electricity. To understand the relationships among knowledge building, epistemic understanding, and conceptual understanding, correlation and hierarchical regression analysis were conducted. It was found that knowledge building environment and post epistemic understanding significantly contributed to students' post conceptual understanding. Chan and Lam (2010) examined role of knowledge-building in facilitating epistemic and conceptual growth, the current study extended it to integrate epistemic dimensions in the knowledge-building design and suggested the importance of helping students to become aware of and resolve their epistemic doubt by scaffolding their collective epistemic inquiry and reflection for epistemic change.

In this study, the key design involved scaffolding children's work as "communities of scientists" and knowledge builders pursuing inquiry into problems, constructing explanations, using authoritative sources of information, improving their ideas and pursing for collective inquiry and new knowledge. While most studies on knowledge building have examined elementary children's scientific understanding, we explicitly focused on helping students to have epistemic reflection. We attempted to help students to reflect on the nature and process of science focusing on theory building as they experienced their own knowledge building processes when they studied electricity. Qualitative analysis of students' interview reflection suggested how the designed environment may have helped students understand science as a collective theory building process, and subsequently helped them revise and improve their theories and explanation.

Although scientific inquiry is much emphasized, students often think of science as lists of activities and skills rather than an idea-driven and theory-building process for creation of new knowledge. This study explored a design that helped students to *mirror* their understanding of science with their engagement in knowledge building inquiry, that merits further investigation. As well, the study extended the line of inquiry on epistemic understanding of science that focused on the role of idea in science, and expanded on social and community processes of scientific progress; these findings suggest possible research direction on examining social aspects

of epistemic cognition. This study is an ongoing study and further analyses of process dynamics would be undertaken to examine the nature of collaborative discourse on Knowledge Forum, and to understand the epistemic and conceptual change mechanism (e.g., the role of epistemic doubt) so as to provide a more coherent picture about the relationship among knowledge building, epistemic cognition, and conceptual understanding.

References

Bendixen, L. D. (2002). A process model of epistemic belief change. In B. K. Hofer & P. R. Pintrich. (Eds.), *Personal epistemology: the psychology of beliefs about knowledge and knowing* (pp.191-208). Mahwah, NJ*:* Lawrence Erlbaum Associates.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works: a study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*(5), 514-529.

Carey, S., & Smith, C. (1993). On Understanding the Nature of Scientific Knowledge. *Educational Psychologist*(28), 235–251.

Chan, C. K. K., & Lam, I. C. K. (2010). Conceptual change and epistemic growth through reflective assessment in computer-supported knowledge building. *Paper in Proceedings of the 9th International Conference of the Learning Sciences 2010.* Chicago, IL.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175-218.

Chuy, M., Scardamalia, M., Bereiter, C., Prinsen, F., Resendes, M., Messina, R., Hunsburger, W., Teplovs, C., & Chow, A. (2010). Understanding the nature of science and scientific progress: A theory-building approach. *Canadian Journal of Learning and Technology, 36*(1).

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology, 29*(2), 186-204.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal epistemology: the psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lederman, J. S., & Ko, E. (2004). Views of scientific inquiry, Form E. *Unpublished paper. Illinois Institute of Technology, Chicago, IL.*

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. e. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*(6), 497-521.

Lee, E. Y. C., Chan, C. K. K., & Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning, 1*(1), 57-87.

Mason, L. (2000). Role of anomalous data and epistemological beliefs in middle school students' theory change about two controversial topics. *European Journal of Psychology of Education, 15*(3), 329-346.

Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary-school students learning science concepts from text. *Journal of Educational Psychology, 87*(2), 282-292.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634-656.

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge In B. Smith (Ed.), *Liberal Education in a Knowledge Society* (pp. 67-98). Chicago: Open Court.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction, 18*(3), 349-422.

van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences, 16*(2), 175-220.

Acknowledgement

The preparation of this paper was supported by a General Research Fund grant from the Research Grants Council of Hong Kong (Grant #HKU 740809H).

Becoming an Activist-Mathematician in an Age of Austerity

Indigo Esmonde, Joe Curnow, Dominique Riviere, University of Toronto, Toronto, ON, Canada
Email: indigo.esmonde@utoronto.ca, joe.curnow@mail.utoronto.ca, dominique.riviere@utoronto.ca

Abstract: In the wake of the global economic crises of the 21st century, governments worldwide have implemented austerity policies, involving dramatic reductions in expenditures. These economic policies provide context for mathematization, as ordinary people grapple with the ways these policies affect their daily lives. In this article, we present an analysis of interviews with members of a Toronto-based activist network that fought against the City's proposed budget cuts in 2011 and 2012. We found that the context of activism spawned a wide variety of mathematics problems related to the activists' overarching goals: understanding the City's use of mathematics, creating their own mathematics, storytelling to the public, and behind-the-scenes tactics. We highlight the ways mathematical activity was always ideological, always accomplished through a division of labor within the group, and always strategically selected. We close by considering processes by which activist mathematicians can be better supported by formal and informal education.

Mathematical Activism in an Age of Austerity

In the wake of a controversial January 2012 Toronto City Council meeting to amend and approved the proposed budget (including hundreds of millions of dollars worth of cuts), both supporters and opponents of the budget cuts announced their victory. According to our informant Maria (all names of activists are pseudonyms), Mayor Rob Ford declared his victory because for the first time in many years, there had been no increase to the City's operating budget. Activists and other opponents to the mayor's cuts celebrated that they had defeated $80 million in proposed cuts to public services and community programs.

Both sides used mathematics to justify their perspective, yet their statements were diametrically opposed. Which side was correct? We argue that both were correct, and this example highlights several interesting aspects of mathematics outside of school: *what* is computed is just as important as *how* it is computed, and the choice of mathematical problems and methods has an ideological component. The budget conflict in Toronto in 2011 and 2012 is but one example of contemporary debates about economic policy, debates that are taking place at the local, provincial, national, and international levels in the wake of the global economic downturn in the 21st century. Governments at all levels have introduced austerity policies, in which they dramatically reduce spending, resulting in cuts to services that disproportionately and negatively affect poor and working class people. These policies affect everyone, however; therefore everyone has a stake in understanding the policies and in making their voices heard either for or against (and often, somewhere in between) the policies that their government representatives put forward. Because these policies have a mathematical component, people who represent all sides of the debate around this issue are called on to become political mathematicians.

In the fall of 2011, when international Occupy movements were talking about "the 1%" and Toronto activists were holding mass public meetings to protest City budget cuts, we began to investigate how activists become mathematicians. The public nature of activism makes activists a particularly accessible group of people to study, if we are interested in how people are mathematizing their understanding of contemporary economic policy. When faced with complex equations rooted in peoples' real lives, like how many people will be displaced from a homeless shelter or how many jobs will be eliminated through restructuring, these activists have necessarily become mathematicians. Their work is not only the public protest that people may expect, but also includes mathematical negotiation, interpretation, and mobilization in order to contest the mathematics circulated by those in favour of austerity, and to replace this mathematics with their own mathematically-based narratives of the impacts of austerity on their communities.

Our analysis examines one case study of activist work to understand the ongoing work of practicing and becoming a mathematician and an activist in an age of austerity. The major issues addressed in our study include the relationship between identity, becoming, and mathematics in activist networks, as well as an analysis of everyday mathematics as it is employed and contested in activist work. This study sheds light on the politicized nature of mathematics learning and practice and offers new contexts for studying mathematical learning in community settings.

In the analysis that follows, we introduce a theoretical framework centred on sociocultural and situated theories of learning. After explaining our methodological approach, we then turn our attention to the Toronto Stop the Cuts network (StC), briefly describing their work before analysing the ways activists in the network talked about and used mathematics in order to contest the austerity agenda of Toronto Mayor Rob Ford. We focus particular attention on activists' overarching goals, and the ways that mathematics intersected with other forms of knowledge as they worked towards these goals. We conclude with a discussion of how mathematics

for social justice was accomplished collectively, and consider the implications for supporting and enabling people to better learn, teach, and deploy mathematical strategies in their struggles for equity and social justice.

Learning as Becoming

Learning is a social process that happens when people engage in collective practices together (Saxe & Esmonde, 2005). Collective practices are "semi-stable, socially organized activities in which individuals participate and communicate with one another" (p. 176). Over time, as people work together to get things done, regularities begin to emerge in the kind of problems that are framed, the resources that are used to solve these problems, and the roles that people take on. Of course, each new problem or situation is an opportunity to begin anew, to adapt and change the resources one uses, or to take on a new role.

In sociocultural and situated frameworks for understanding learning, learning can be understood as a process of becoming, both individual and collective, that is always understood in relation to the practice in which one learns. For example, children who were competent candy-sellers on the streets of a Brazilian city, were mostly unable to complete school mathematics tasks, and successful schoolchildren were unable to complete the tasks required for candy-selling (Saxe, 1991). As Saxe points out, the goals that are constructed in these two contexts differ. In candy-selling, the main goal is to make a profit and to sell a lot of candy. In school, the goal is usually to solve problems in a teacher-approved way. The resources also differ, with a much wider variety of strategies, tools, and other people available in out-of-school contexts.

Similar to Saxe's findings, research in mathematics outside-of-school has predominantly found that people are quite successful in the non-school mathematical tasks that they encounter on a daily basis, in the workplace, in the home, and in their hobbies (Esmonde et al., 2013). Again, this competence should be understood in context: it is not that people are smarter outside of school. Rather, non-school contexts afford many more resources than classrooms typically do, and people have much more freedom to organize their activities in ways that lead to success. Most striking, for our purposes, is the collaborative nature of many out-of-school activities. Like social justice activism, carpet-laying (Masingila, 1994), dairy-work (Scribner, 1984), home improvements (Esmonde et al., 2013), and video games (Stevens, Satwicz, & McCarthy, 2008), all allow people to help one another and ask for help, and even allow them to extend the division of labor beyond their immediate circle by using a variety of resources, including seeking expert help (via the internet, visits to hardware stores, etc.).

Research examples outside of school make particularly visible the role of *values* in mathematical practices. Although mathematics is often seen as a positivist science, mathematical models of the real world are based on a worldview, and therefore on a set of values (Frankenstein & Powell, 1994). In home improvement, one may balance aesthetic values with a value for cutting costs. In game play, one may search for a particularly elegant strategy. School mathematics also encodes a set of values that are embedded not only in the textbook examples of mathematics, but in the ways students and teacher are expected to interact together. In contexts in which these values are brought to the fore (such as, contexts in which ideology is explicitly discussed and debated), the relationship of one's values to one's mathematics will be more clearly visible.

Drawing from this theoretical framework, in this study we investigate a context for mathematics practices in which the work is collaborative, mathematical work is embedded with other forms of content knowledge, and the values underlying mathematics are made explicit. In an age of austerity politics, social justice activism is a hub of value-laden mathematical work.

Methods

Toronto Stop the Cuts (StC) [http://www.torontostopthecuts.com/] was a network of community organizations, concerned individuals and neighbourhood committees who advocated to stop the budget cuts being proposed by Mayor Rob Ford in Toronto. Formed in 2011 as a response to the cuts to city services outlined in the proposed City budget for 2012, StC organized meetings in various neighbourhoods across Toronto and urged residents to take actions against the cuts to services such as transit, housing, daycare, and shelters. These actions included meetings, marches, dinners, councillor visits, lemonade stands, etc. StC also advocated through organizing online campaigns such as citywide petitions. Overall, StC was committed to forming relationships with allies in order to fight austerity at all levels. In particular, the network had three main demands for the City's Mayor and councillors: 1) Stop the budget cuts to public services and, relatedly, to stop the attacks on public sector workers in an effort to advance a privatization agenda; 2) Expand city services for all, because StC argued that public services in the city were inadequate even before the proposed budget cuts for 2012; and 3) Increase the corporate tax rate, and reduce spending on the City's police services. Implicit in these demands was a critique of austerity policies that "bail out" corporations while sacrificing public services, and a fear that increases in the police service's budget would negatively affect, and disproportionately so, the people who would be most affected by the other budget cuts (the city's poor and working class, as well as racialized and undocumented people).

The StC network reflected the diversity of the city of Toronto, including diversity as considered in terms of race/ethnicity, gender, sexuality, immigration status, neighborhoods in Toronto, socioeconomic status, age, linguistic background, and more. StC operated as a network with a number of subgroups. The primary focus of organizing happened at the neighborhood level. StC started several neighborhood groups and encouraged anyone in the City to start their own neighborhood group and join the network. The only requirement was that all members had to agreed with StC's three main demands as described above. Each neighborhood group was expected to meet regularly and organize actions that made sense within their own contexts. However, these subgroups were expected to bring their decisions to the "network meeting" for discussion and approval. The network meetings consisted of regular meetings with representatives from all neighborhood groups and committees. At these meetings, neighborhood groups updated one another on their news, jointly decided on policy, collaborated on larger actions and campaigns, and made requests for help. There were several committees that were created to support the work of the neighborhood groups and the network as a whole. These included the research committee, tasked with doing research to answer questions that arose at the neighborhood level (e.g., creating reports for each neighborhood committee about the demographics and key issues in their neighborhoods), and the media committee, which was tasked with creating press releases and monitoring the media for news about the budget process.

For our research, two of us joined StC as participant observers for four months, from October 2011 to February 2012. We were members of the Research Committee and helped field research requests from the various neighborhood groups. We also collected data at various public meetings and protests. Our dual role as researchers and participants was clear to all StC members with whom we interacted. Our participation probably had a positive influence on our ability to recruit interview participants (described below), and also informed our data analysis because we had sufficient background knowledge to be able to interpret the interviews.

Data and Interview Questions

This paper is based on an analysis of interviews with seven activists from the StC network. These activists were all engaged regularly at network meetings, in addition to their activity within their neighbourhood group or subcommittee. Our interview pool was small, but generally reflected the diversity of the StC network and the city as a whole, with the exception that all of our interview participants had post-secondary degrees. We had recruited our participants by inviting anyone active with StC to conduct an interview on the use of mathematics in activism. Informally, we were told that many people were reluctant to agree to an interview because they did not feel qualified to discuss mathematics. Since our analysis in this paper focuses on elaborating the range of mathematical practices of StC, we feel that our interview pool was adequate, but readers should remember that we were informed by a set of participants who felt more comfortable with mathematics than perhaps the average activist did.

Each activist was invited to an hour-long, semi-structured interview. The interviews were video-recorded and stored on a secure data drive at the University of Toronto. The interview questions were informed by our theoretical framework in that our goal was to obtain a broad understanding of the collective practices of activism, and the specific role of mathematics as one aspect of the broader practice. The questions focused on elaborating the context of StC's activism by asking them about the goals and mission of StC, the City's use of mathematics, and StC's use of mathematics. We asked activists to be specific about their own roles and responsibilities, especially in relation to mathematics, and we asked them what helped, and what hindered, people in doing, using and learning mathematics towards social justice.

Research Questions and Data Analysis

There was one major research question for this study: *How is mathematics used outside of schools to analyze, educate about, and take action towards broad social justice goals?* Within this question, we were interested in both the types of mathematical tasks that were taken on, as well as the process by which these tasks were accomplished. However, as noted in the introduction to this article, we do not believe that it is useful to analyze mathematical tasks separate from the context in which they arose. Thus, we were also interested in understanding the complex forms of knowledge that were woven together in the work of Stop the Cuts activists.

To investigate this research question, we began with creating a content log for our seven interviews. These content logs recorded a basic summary, minute-by-minute, of what was discussed in each interview. Following this, we began with a very basic round of coding, knowing that we were interested in how the activists identified the mathematics that was used ("kinds of math") and the ideological nature of mathematics ("ideology"). As the analysis progressed, we created a listing of all the different mathematical tasks that were named by the activists. This list was separated into themed categories, based on the overarching goals to which mathematics was being used. (It is important here to note that due to the nature of the interview data, we were unable to gain insight into the details of the mathematical tasks or the process by which they were accomplished. Such an analysis would require a very different set of methods. Thus, rather than analysing mathematical *reasoning*, here we present an analysis of the activists' stories about the ways they used

mathematics, as well as its place within the broader collective practice.) After completing the coding of various kinds of math, we considered what the activists' stories about mathematics could tell us about some of the central aspects of our theoretical framework: resources, division of labor, and ideology/values.

Our results, as presented here, rely on an analysis of the interviews alone, but we were able to draw from our background knowledge and participation in the group to help interpret some of the stories we were told. At times in our presentation of findings below, we mention specific documents or websites that were used or created by StC. These artifacts were all mentioned in the interviews, and our discussion of them is limited to what we were told in interviews. Below, we link to specific documents or sites to make them accessible to readers interested in the specifics of activist mathematics, but we did not do a formal document analysis for this paper.

What Is Activist Mathematics?

We identified four overarching mathematical goals, which form the basis for the bulk of the analysis in the paper: understanding the City's use of mathematics, countering with their own mathematics, storytelling to the public, and behind-the-scenes tactics. All of the interview participants were able to describe many instances of the use of mathematics to achieve the four activist goals. We will discuss each of these themes in turn, but we acknowledge that activities within each theme were interconnected; for example, understanding the city's use of mathematics informed activist strategies for mathematics, as well as the storytelling that they engaged in for the public.

Understanding the City's Use of Mathematics and its Implications of the Budget Cuts

In order to fight the budget cuts, first the activists needed to understand precisely what the cuts were, and further, they wanted to go beyond a surface understanding of the cuts to understand how the cuts would affect their communities. Activists expressed especial concern about the effect of the cuts on communities that were already marginalized, including racialized people, undocumented people, homeless people, and people living in poverty.

One of the most foundational tasks for StC was to deconstruct the City's reports, including the Core Services Review (City of Toronto, 2011) that had been produced by external consultants in order to find the 'gravy'[1] in City services. These reports (which can be found on the City's website at http://www.toronto.ca/torontoservicereview/results.htm) were filled with charts, tables, and mathematical arguments. To further understand the impact of the cuts, StC members created maps (thus using spatial and geometric reasoning) to display which services were likely to be affected by the cuts, in which neighborhoods. An example of such a map was displayed on the StC website (http://www.torontostopthecuts.com/january-10-%e2%80%93-mapping-the-cuts-part-ii/) and was used to argue that the City's high-poverty neighborhoods would be more directly impacted than wealthier neighborhoods. The map included locations of cuts to public housing, libraries, homeless shelters, childcare centres, public recreation centres, and long-term care homes. The map was created by another Toronto-based advocacy group, Social Planning Toronto (http://www.socialplanningtoronto.org/). StC borrowed the map and displayed it on their website.

Some of these numbers required further analysis. For example, the activists had a special focus on childcare, and were able to draw from information contained in City websites to determine the number of existing childcare spots available in each neighborhood, and the number of spots likely to be cut. They could then consider the impact of these cuts by considering the number of people who would be unable to work due to inadequate childcare. For libraries, activists identified each of the libraries that would face cuts and were able to pinpoint how many hours of service would be cut, or the numbers of staff, and from there, were able to discuss the impact on library services. They drew on a proposal for cuts to the City's public transit network (Toronto Transit Commission, TTC) to alert City residents to which bus routes would have longer wait times. For cuts to homeless shelters, they could describe how many more people would be "out on the street" than previous. In all of these instances, StC members were required to make sense of pre-existing mathematical calculations, and organize them in ways such that their real-life impact could be easily understood.

Countering With Their Own Mathematics

One major mathematical project was the creation of the activists' People's Poll (with results reported on the StC site (http://www.torontostopthecuts.com/peoples-poll-results/). This poll was created for multiple reasons, one being that the activists felt that the City's public consultation process was highly skewed. Aziz told us that "We didn't see uh, these, city-run, uh, town halls, which produced their own statistics, as legitimate by any means." He said that within StC there was wide agreement that "we need our own statistics." The activists were concerned about several aspects of the City's polling. In interviews, several participants described how group members had disagreed with the method of recruitment, because they felt that many City residents were unaware of the polls that were conducted online and at Town Hall meetings around the City. In addition to recruiting online, StC's major strategy was to go to public places to recruit participants for the People's Poll. A

second critique of the City's data-gathering was that StC felt that the City's poll questions were biased towards cuts. Stephanie pointed out that the City was asking people *which* services should be cut, rather than *whether* services should be cut.

It should be noted that there was some disagreement within StC about the use of the People's Poll. Some felt that since their poll "would never pass muster in terms of how surveys should be done" (Maria) and acknowledged that it would not be "statistically reliable" (Maria). Still, they decided to conduct the poll so that they would have some control over the data collection and reporting process. In addition, they felt it was a valuable organizing tool because it allowed StC members – especially new members – to go out into communities and talk with people about the issues. Thus, this mathematical task also served non-mathematical purposes.

The interviews did not contain many other specific examples of StC activists countering with their own mathematics, although several interview participants referred to this general process. For example, Aziz told us about some analysis StC had done on library use. He said, "our issue isn't so much with the use of math. Our issue is how are you using it and what are you quantifying." The City had decided in advance that it would cut hours at selected libraries. In choosing libraries to cut, our interview participants said that the City had looked mainly at circulation (how many books are checked out). StC felt that this was not a good measure of the use of libraries; instead, they looked at the programs that libraries supported, and the number of people who made use of the various resources at the library – internet, reading rooms, children's programs, and more.

Maria reported on a mathematical task that StC had decided *not* to do, for ideological reasons. She told us that some members of StC had wanted to create an alternative budget, to recommend how the city should use its funds. Others felt that this process would "pit vulnerable groups against each other" (Maria). This strategy went against StC's three basic demands and was ultimately rejected. As Peter pointed out, liberals who did not oppose capitalism and were simply trying to shift the budget slightly would need to argue that their proposals "add up" mathematically (i.e., that there is enough funding for their proposals). As a radical anti-capitalist activist, Peter did not feel that StC was obliged to argue that their three demands were economically feasible. As he put it, StC was "not concerned with the wellbeing of their [capitalist] system."

Storytelling to the Public

StC activists told us that numbers and mathematical arguments were an important part of their campaign to convince the public of the rightness of their cause. "The use of statistics kind of carries this very factual, very objective kind of truth to it," said Aziz, and this statement about the rhetorical value of mathematics was supported by several other interview participants. On the other hand, four out of the seven interview participants mentioned that they believed that many people (activists included, but not limited to activist circles) were afraid of mathematics and didn't feel qualified to engage in argumentation about the big numbers that were used in the budget. Thus, although StC wanted to use numbers in flyers, speeches, and press releases, they were cautious to make those numbers relatable to people (although they admitted to, at times, using big numbers as a scare tactic, just as they felt the City was doing).

"Organizing is storytelling," Ahmed told us, as he emphasized the importance of constructing stories that people can relate to. The People's Poll was one tool that StC used in this storytelling, to convince the public and the City government that vast numbers of people were opposed to budget cuts and austerity measures. The results of the People's poll (and, in fact, the results of the City's polling, flawed as they believed it to be) gave StC a kind of "numerical legitimacy" (Ahmed) to show that overwhelmingly, City residents rejected the proposed cuts.

Other examples of mathematics used in storytelling included key numbers that were included on flyers or press releases. For example, when StC discussed cuts to public transit or a proposed expansion of TTC services, they sometimes included figures about ridership, or the number of people who have access to the TTC, speed, or cost. They would often place their own mathematical arguments side by side with the City's, to argue that the City's plan did not make mathematical sense. For example, when the mayor argued that the City would have a very large deficit (approximately $700 million), activists pointed out that he had cut the vehicle registration tax ($60 per car) and the land transfer tax (a tax on real estate purchases), and that if he had not made those cuts, the supposed deficit would be significantly smaller (both Ahmed and Stephanie discussed this storytelling narrative).

The activists chose their numbers carefully for maximum effect. For example, Ahmed argued that large numbers are very difficult for people to understand, so instead of presenting a large number like, for example, 3000 people, StC would say something like "twenty-seven subway cars" (Ahmed came up with this number as an example during his interview; thus, it may not reflect an actual calculation that StC made, but is an example of the type of calculation). They argued that numbers like these were easier for people to relate to. However, in the spirit of storytelling and relationship-building, they sometimes chose not to represent numbers. Ahmed said that early on in the organizing, they tended to use more data and numbers in their arguments to the public. Later on, when they had built relationships with individuals in different communities who would be impacted by the

cuts, they tended to report more personal stories. For example, instead of reporting on the number of people who would be impacted by cuts to WheelTrans (a public transit service for people who use wheelchairs), they might create a video with one person telling their story and describing how they would be affected if they could not access WheelTrans anymore.

Behind-the-Scenes Tactics and Strategizing

The work of organizing – of reaching out to as many people as possible through flyers, news and social media, public actions and protests – requires many skills, including some mathematical skills. Since StC activists primarily focused their organizing efforts at the neighborhood level, they gathered publicly available information about the demographics of different neighborhoods so that they could tailor their organizing efforts to these neighborhoods.

When creating banners, flyers, and other graphic images, they used mathematics as they measured distances and centered text. When planning a march, they estimated the number of copies of flyers they should make based on the numbers of StC activists who could go out and flyer in the various neighborhoods. They estimated the number of participants they would have (using information from social media and other sources) and based their recruitment efforts on achieving targeted numbers for various actions. For example, they used a rough estimate that about half the people who rsvp'd for an event on Facebook would actually show up. At a march, they would count the number of participants by coming up with an estimate of the number of people who would fill a certain amount of space (e.g., a ten by ten foot square), and then estimating how many ten foot squares were filled. Or, they found estimates on the Internet of the number of people who could fill the street in a standard city block, two lanes wide, and then use this information to estimate turnout. When they timed events that included marches, they estimated the amount of time it would take to march from place to place, especially if there were several stops with speakers or banner drops.

They also collected data on their organizing efforts and analyzed them to see if they were reaching a broad demographic of the City. For example, they used Google analytics to see when people were accessing the Stop the Cuts website, and they collected demographic information from signatories to a Declaration that they had produced, to see who was joining the StC cause.

They used data to target their efforts in ways that they felt could be more fruitful. For example, using data about councilor voting patterns, they targeted city councilors in what was known as the 'mushy middle': rather than working to convince councilors who always voted for or against the mayor's policies, they focused on councilors who were more variable in their allegiance. They also used data regarding city resident voting patterns to figure out which neighborhoods were more aligned with the mayor, and which were less supportive.

Finally, within StC when there were conflicts, with some activists pushing for more radical or confrontational tactics, and others arguing for more mainstream methods. Some of the activists drew from data to argue that StC had not yet been successful in its aims, and an escalation of tactics was necessary.

Mathematics in Relation to Other Activist Tools

Given our theoretical framework, it is important to consider the *process* of activist mathematics, not as a series of disconnected cognitive tasks, but as part and parcel of the work of activist groups. In this section, we discuss three findings about the process of activist mathematics that align with, and extend, our theoretical framework. These findings include: the necessity of appropriate content knowledge in order to mathematize inequality; the division of labor within the group; and the intentionality of activist mathematics.

As Aziz pointed out, StC activists often disagreed with the City's mathematics because they differed in "how you are using [mathematics] and what you are quantifying." For example, with the library cuts, the City used primarily circulation numbers, whereas StC argued that the number of people who used the library, for many reasons beyond just taking out books, should be taken into account. In order to decide how to mathematize a real-life situation, deep content knowledge was necessary in order to make sense of how a given political and economic decision would affect people's everyday lives. StC's disagreements with the City highlight that mathematical modeling of the world is *always* ideological. The ideological nature of mathematics was revealed throughout the data reported above. StC's deconstruction of the City's math aimed to reveal the underlying ideology of the cuts, their own mathematics was explicitly based on their ideological perspective (e.g., not being concerned about "the well-being of [the capitalist] system"), they selected numbers (or decided against numbers) to make their stories to the public as convincing as possible, and their behind-the-scenes strategies were based on an understanding of democratic ideologies in which getting as many people involved as possible, to show "numerical legitimacy," was paramount.

In a related point, StC members often expressed much more certainty about their ideological standpoint than their mathematical models. As Alex explained to us, in StC there were "more people who could explain Marxism to you than could explain fractions." Some mathematical tasks could be performed by any newcomer to the group: collecting data for the People's Poll was considered an entry level task that could help newcomers learn about the issues and build connections to the broader community. Other tasks, such as reading and

summarizing government reports, or collecting demographic information about the various neighborhoods, fell primarily to the Research Committee. StC was an example of a community of practice in which various specializations emerged; no one member of the group would have been able to accomplish all that the group could accomplish together. Mathematics was one, but certainly not the only, aspect of activist work that was specialized to a subgroup. In this case, the division of labor extended beyond StC into other groups and individuals who were pursuing similar goals. Through the Internet, news stories, and public forums and events, StC had access to the mathematical activist work of other organizations.

Finally, this study's findings differ from most out-of-school mathematics research in one significant way: the intentionality of the mathematics. In many studies of people's participation in out-of-school activities, mathematics emerged as an invisible aspect of their work. People did not always recognize that the tasks they accomplished every day were mathematical. StC resembled these studies in one way: many of our interview participants told us that the group never discussed the question of whether they should or shouldn't use mathematics. Rather, they discussed specific tactics: Should we conduct a poll? How might we analyze it? Would it pass muster? However, despite the lack of the label "mathematics," the discussions and actions of the StC group showed the intentionality of their engagement in mathematics. All of our interview participants told us that the City was using mathematics to intimidate the general public by using big numbers and doomsday language about deficits and inevitable cuts. With mathematics used as a weapon against them, all of the activist interviewees told us that it was critical for them to be able to understand, and to counter, the City's use of mathematics. This suggests that community activism may actually be fertile ground to support people in understanding and using mathematics more intentionally and towards liberatory ends.

Implications: How Can We Foster Activist Mathematicians?

The StC campaign against austerity agendas in Toronto was partially successful. The group's role was instrumental in eventually defeating some of the budget cuts and preserving most City services for the 2012 budget. While the findings of this study, related to the grassroots mobilization of mathematics for social justice, are interesting in their own right as an example of informal mathematical engagement, we believe these findings also raise some questions about the role of formal education in supporting progressive social change.

If we consider Alex's argument that more activists could explain Marxism than fractions, several questions are raised. First, how can mathematics education support people in learning, doing and using mathematics for social justice? In recent years, many educators and researchers have brought social justice into mathematics classrooms, with some success (Enyedy & Mukhopadhyay, 2007; Gutstein, 2006; Gutstein & Peterson, 2013; Turner, Gutiérrez, Simic-Muller, & Díez-Palomar, 2009). We would like to see more studies, like Gutstein (2006) in which mathematics students describe how these experiences supported their mathematically-engaged activism in their lives outside of school. As an anonymous reviewer suggested, the four themes we presented in our findings (understanding government use of mathematics, countering with their own mathematics, storytelling to the public, and behind-the-scenes tactics) might serve as a useful heuristic for organizing mathematics teaching for social justice.

Second, the flip side of Alex's argument suggests the need for social and political education that integrates mathematics as a lens through which to see the world. Why is it that so many StC activists were comfortable with the complexities of Marxism but frightened of the complexities of the City budget documents? Perhaps mathematics education researchers could venture into contexts devoted to social and political education, to develop an understanding of how mathematics is used (or possibly pushed to the background) in such contexts.

Finally, many of those responsible for mathematical aspects of StC's activism were from middle-class backgrounds with post-secondary degrees. How can educational contexts that are not limited to K-16 schools support the integration of mathematics with all of the complex knowledge necessary for imagining and working towards a better world?

Future studies in this area could address all of these issues: the ways in which mathematics education supports the mathematization of global and local economic policy; the ways in which social sciences or political education support people in drawing mathematics in to their understandings, and the ways people learn to mathematize the world through participating in out-of-school community organizing. We are particularly interested in detailed analyses, to supplement our broad view of the role of mathematics, to uncover tools, resources, and divisions of labor, which are effective at broadening people's participation in political mathematics.

Endnotes

(1) Ford had promised that if elected, he would end the "gravy train" of City services and employment. The phrase refers to a situation in which a person can earn a lot of money with very little effort.

References

City of Toronto. (2011). *Core Services Review*. (Prepared by KPMG). Toronto, ON: Canada. http://www.toronto.ca/legdocs/mmis/2011/ex/bgrd/backgroundfile-40702.pdf

Enyedy, N., & Mukhopadhyay, S. (2007). They don't show nothing I didn't know: Emergent tensions between culturally relevant pedagogy and mathematics pedagogy. *Journal of the Learning Sciences, 16*(2), 139-174.

Esmonde, I., Blair, K. P., Goldman, S., Martin, L., Jimenez, O., Pea, R. (2013). Math I am: What we learn from stories that people tell about math in their lives. In B. Bevan, P. Bell, R. Stevens (Eds.), *Learning in out of school time.* (pp. 7-28). New York: Springer.

Frankenstein, M., & Powell, A. B. (1994). Toward liberatory mathematics: Paulo Freire's epistemology and ethnomathematics. In P. L. McLaren, & C. Lankshear (Eds.), *Politics of liberation: Paths from Freire* (pp. 74-99). London: Routledge.

Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice*. New York: Routledge.

Gutstein, E., & Peterson, B. (Eds.) (2013). *Rethinking Mathematics: Teaching Social Justice by the Numbers*. Milwaukee, WI: Rethinking Schools.

Masingila, J. O. (1994). Mathematics practice in carpet laying. *Anthropology & Education Quarterly, 25*(4), 430-462.

Saxe, G. B. (1991). *Culture and cognitive development: Studies in mathematical understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Saxe, G. B., & Esmonde, I. (2005). Studying cognition in flux: A historical treatment of Fu in the shifting structure of Oksapmin mathematics. *Mind, Culture, and Activity, 12*(3&4), 171-225.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff, & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 9-40). Cambridge, Massachusetts: Harvard University Press.

Stevens, R., Satwicz, T., & McCarthy, L. (2008). In-game, in-room, in-world: Reconnecting video game play to the rest of kids' lives. In K. Salen (Ed.), *The ecology of games: Connecting youth, games, and learning* (pp. 41-66). Cambridge, MA: MIT Press. doi: 10.1162/dmal.9780262693646.041

Turner, E. E., Gutiérrez, M. V., Simic-Muller, K., & Díez-Palomar, J. (2009). "Everything is math in the whole world": Integrating critical and community knowledge in authentic mathematical investigations with elementary Latina/o students. *Mathematical Thinking and Learning: An International Journal, 11*(3), 136-157.

Acknowledgments

This research was funded by the Social Sciences and Humanities Research Council of the Government of Canada. We would like to thank the members of StC for allowing us some insight into their important activist work. We would also like to thank Scott McDonald, Jennifer Langer-Osuna, Niral Shah, GLITTER (Group for the study of Learning, Identity, and Teaching Towards Equitable Relations, including at that time Lesley Dookie, Allison Ritchie, and Miwa Takeuchi), and anonymous reviewers for reading an early draft and providing feedback.

Measuring Affective Experience in the Midst of STEM Learning

Jayson Nissen, Jonathan Shemwell, University of Maine, 120 Bennett Hall, Orono, ME 04469,
Email: jayson.nissen@maine.edu, jonathan.shemwell@maine.edu

Abstract: The Experience Sampling Method, an at-the-moment survey technique, was used to measure university students' affective experiences within school and in their daily lives on four variables: activation, self-efficacy, motivation, and stress. Affect was compared for school vs. non-school, and within school, STEM coursework vs. non-STEM coursework. Within STEM, affect for a focal physics course was compared to affect for all other STEM courses. School was experienced with higher stress, lower intrinsic motivation, and lower self-efficacy than non-school. STEM and non-STEM courses were not experienced differently, but the physics course was experienced with higher stress and lower self-efficacy than other STEM courses. The results suggest that, broadly, university coursework may undermine intrinsic motivation and that the negative impact occurs in the midst of instruction. More tentatively, the process of engaging with challenging STEM content, such as that of the physics course, may tend to increase stress and undermine self-efficacy.

Introduction

Increasing the number and diversity of students who enter and remain in the STEM education pipeline is an important goal in the United States and in many developed countries (National Commission on Mathematics and Science Teaching for the 21st Century, 2000; Osborne & Dillon, 2008). Attainment of this goal will naturally depend on positive learning outcomes for diverse students in K-16 STEM education. One important and often overlooked class of outcome is positive affect toward STEM domains. Two examples of affect are feelings of self-efficacy and motivation. Declining STEM enrollments and increasingly negative attitudes toward STEM suggest that positive affect is in short supply within STEM instruction (Osborne, Simon, & Collins, 2003; Semsar, Knight, Birol & Smith 2011). Understanding the nature and sources of affect within the current STEM education system is therefore an important research goal.

Most studies of affect in STEM education have utilized survey techniques in which students report their general attitudes and beliefs for an overall experience, such as at the end of a course. Many studies have related affective variables measured by surveys to important outcomes such as achievement. For example survey-based studies in STEM have measured the possible mediating effect on student achievement of self-efficacy (Lee, 2009; Marra and Bogue, 2009), interest (Koller, Baumert & Schnabel, 2001), and motivation (Singh, Granville & Dika, 2002; Mujtaba and Reiss, 2013). Mujtaba and Reiss (2013) analyzed student end of course responses to a survey to show that the lower representation of women in physics was related to differences between male and female secondary student's affective experiences in physics. Lee (2009) used survey responses to investigate the factorial relationship of three affective traits within math, self-concept, self-efficacy, and anxiety, and how their relationship to student achievement varied between countries. Luzzo, Hasper, Albert, Bibby and Martinelli (1999) used pre/post surveys to investigate the impact of a self-efficacy-enhancing intervention on student's math and science self-efficacy and career interests. In a mixed methods approach Girod, Rau and Shepige (2002) used both surveys and interview case studies to demonstrate that elementary school students had a higher quality of experience in a course focused on aesthetic understanding as opposed to one focused on conceptual understanding.

A challenge in studying the impact of instruction on affect is that it is difficult to measure affective response in the midst of the experience. Holleran, Whitehead, Schmader and Mehl (2010) used analysis of random audio samples from day to day life to overcome this difficulty. They showed how stereotype threat affected how female STEM faculty interacted with other female faculty as opposed to male faculty. Stephens (2012) reproduced the classroom experience in the laboratory in order to measure stress caused by the mismatch between the culture of college and that of first generation students on students. Stephens measured stress by periodically collecting saliva (for cortisol analysis) while students gave speeches they had written.

Our approach to measuring affective response within instruction is to adapt a technique called the Experience Sampling Method (ESM) (Hektner, Schmidt & Csikszentmihalyi 2007). We used ESM to measure university students' affective experience in their courses and throughout their day-to-day lives on four distinct categories of affective experience: activation, self-efficacy, motivation, and stress. This method allowed comparing experience between the students' school and non-school activities, and between their STEM and non-STEM courses.

Methods

The purpose of the study was to investigate students' experiences of self-efficacy, stress, motivation, and activation in both non-school and school experiences, and within these school experiences to compare students experience in both STEM and non-STEM courses. A further purpose was to compare affect within a focal, reformed STEM course to affect within other STEM courses.

Context

The study took place at a flagship state university in United States. The focal STEM course was a 15-week large-enrollment calculus-based introductory physics course. The instructor, who had more than 20 years of teaching experience, was on his fifth year of teaching the course. In the previous four years, the instructor had implemented several research based teaching practices that are commonly referred to as interactive engagement (IE) practices (Hake, 1998; Kost-Smith, 2011). These included collaborative problem solving activities in the two one-hour recitation periods each week, and electronic response, i.e. "clicker", questions embedded throughout both one-hour weekly lectures. There was also one two-hour lab each week that was not taught by interactive engagement. Data collected in this "IE Physics" course demonstrated learning outcomes that were similar or superior to those of IE courses at other institutions (Hake, 1998). Specifically, on Force and Motion Conceptual Evaluation (Thornton & Sokoloff, 1998), the IE Physics course yielded normalized gains for the previous four years ranging between 45% and 55%. Grade distributions in the course were approximately evenly distributed with 24% A's, 23% B's, 27% C's and a 26% rate of drop, withdrawal, or failure (DWF). A report obtained from the university's Director of Institutional Research showed that the grading distribution was not different from those of other STEM courses.

Participants

Participants were 33 of the 244 students enrolled in the IE physics course. Half of the participants were female (17). This proportion over-represented the population of females in the course, which was 18%. Participants were recruited by a brief presentation in the IE physics lecture and a follow-up email to those who provided their contact information. Participants had higher final grades than the average for students in the course, 82% versus 76%, but similar distributions in grades awarded. Participants were compensated with extra credit in the IE physics course for the first time they participated, and a check for 50 USD for the second time they participated.

Design

The design was a within-subject comparison of the affective experience within school (i.e., while in class or doing homework) and non-school activities spanning the range of students' day-to-day lives. Experience within school was broken down according to three different types of courses participants were enrolled in: IE Physics, STEM courses other than IE Physics, and non-STEM courses.

Procedures

Data collection occurred during two separate weeklong periods during the third and tenth week of the semester. Twelve of the 33 participants provided ESM data during both data collections. During the data collection weeks, no tests were taken or returned in the IE physics course. Participants completed a one-hour training the week prior to their first week of data collection. Training included a description of the ESM and practice doing the ESM.

During data collection participants in the ESM were semi-randomly signaled using a text message sent to their personal cell phone 5-8 times a day across each one-week period for a total of 50 signals for each week. Five of the signals were scheduled for random times during each of the scheduled IE physics course components (lab, lectures, and recitations) throughout the week. This allowed collecting a large enough number of surveys to provide a representative and diverse sample of experience within the IE physics course setting. Participants received the remaining 45 random text message signal during daily activities: making breakfast, driving, sitting in class, playing sports, etc.

Upon receiving a signal, students would pause what they were doing to fill out very brief at-the-moment survey about their affective experience. Sometimes the signal would occur during a university course or an activity related to a course such as homework; sometimes it would occur during non-school recreational hours or while students were involved in quotidian tasks such as laundry or dishes. Participants' responses began with writing brief statements about what they were doing, what they were thinking, and where they were. Participants would then check a circle for each of 20 Likert-scale affective questions to indicate their affective experience in the activity. Example items drawn from the survey are provided in Figure 1 and described more fully in Table 1. A pilot study showed that completing the survey took 1-3 minutes. Most participants in the pilot study indicated it had little or no impact on their activity.

Some activities (such as driving), prevented participants from providing their responses at the moment of the signal. In accordance with standard ESM procedures, participants were told to complete surveys as soon as possible after the signal and to complete them regardless of how much time had passed. The survey form included a space to indicate the delay between the signal and completion of the survey. Only surveys completed within 15 minutes of the signal were included in the data analysis. This is a standard technique of ESM to help ensure that the responses measure affect as close to the moment of the signal as practicable. Similarly, only participants who completed at least fifteen surveys were included in the data analysis because that is broadly considered to be a minimum threshold for measuring student's average experience when using the ESM (Hektner *et. al.,* 2007). The first author collected the surveys at the end of the data collection period and transcribed the Likert-scale and open-ended responses into a data file containing all responses for each student.

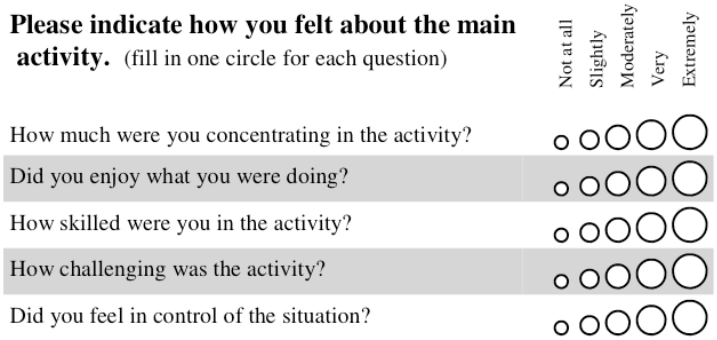


Figure 1: Examples of Likert-scale affective questions used in the ESM data collection.

Construct Definitions and Instrumentation

Table 1 defines each category of effective experience and shows several survey items used to measure it. The survey items were drawn from prior research using ESM (Hektner *et. al.,* 2007). As is done in many ESM studies, item groupings were confirmed in the present study using exploratory factor analysis. As can be seen from Table 1, the measures within the ESM access each of the affective categories broadly and cannot distinguish fine details within these constructs. For instance research on motivation includes concepts such as goal orientation (Belenky and Nokes-Malach, 2012), which are not represented in the present study's formulation. Similar examples of the complexity of the other affect categories defined in Table 1 exist throughout the scientific literature. The present study's coarser measurements were a trade-off for the ability to access multiple categories of affective experience across daily activities for a representative sample of participants.

Self-efficacy was the primary category of interest in designing the research due its obvious implications for attracting and retaining students in STEM. Bandura (1997) defined self-efficacy as the confidence in one's capability to perform the actions necessary to achieve a particular goal. The present study departed from Bandura in that it focused on self-efficacy experiences, as opposed to more stable dispositions or beliefs, by asking how skilled or successful people felt in the activity they were doing (Table 1). However, the present study aligned with Bandura in that the measure was specific to the activity at hand. The self-efficacy measure was based on items with unipolar scales, starting from a zero value and extending to a maximum.

Intrinsic motivation was comprised of enjoyment, excitement, freedom (as opposed to constraint), and importance to future goals. Several of the motivation items had bipolar scales, making the intrinsic motivation measurement either negative (extrinsic) or positive (intrinsic). Intrinsic motivation was measured on the assumption that it is necessary, at some level, for students to opt themselves into a learning experience. Additionally, Deci, Koestner & Ryan (1999) showed the importance of intrinsic motivation by demonstrating the long-term negative impact of extrinsic motivation on persistence and learning.

Stress and activation were utilized as complementary measures to further inform self-efficacy and motivation. For instance, we wanted to see if feelings of low self-efficacy were generally accompanied by stress. Lazarus & Folkman (1984) defined stress as a negative feeling resulting from an individual's perception that they do not have the resources to cope with a perceived situation. Our measure was only loosely aligned with this definition: we asked directly about stress, worry and frustration. All of the stress items utilized unipolar scales. Activation was defined as the level of involvement in the task, consistent with Thayer (1996). In contrast to Thayer's formulation activation was a unipolar measure, which included alertness, attentiveness, and the degree to which the participant was concentrating on the activity.

Table 1: Details for each of the affective categories measured by the ESM

Affect Category	Number of items	Example questions[1]	Definition
Activation	7	How alert did you feel? How attentive did you feel? How much were you concentrating?	The level of involvement in the task at hand.
Self-Efficacy	5	How skilled were you? Did you feel in control? Were you succeeding?	Sense of personal success and capability in accomplishing the task at hand.
Stress	3	How stressed did you feel? How worried did you feel? How frustrated did you feel?	Emotional strain or tension resulting from adverse or very demanding circumstances.
Intrinsic Motivation	4	Did you feel free or constrained? Did you enjoy what you were doing? How important was the activity to your future goals?	The reason for action is drawn from a sense of enjoyment rather than an external reward.

1 Actual questions used in the ESM followed a briefer format.

Methods of Analysis

Completed surveys were coded as either non-school activities or by the type of course associated with the school activity. School activities were then reduced to three approximately equal sized categories: (1) Non-STEM courses which included a diverse range of courses such as English, anthropology, and art; (2) STEM courses excluding IE physics, primarily consisted of chemistry, calculus and introductory engineering courses; and (3) IE physics. Analysis of variance showed that, within both the STEM and non-STEM categories, there were no statistically significant differences in affective responses between courses composing both categories, i.e. calculus and chemistry ($p > 0.20$).

For an activity type, each participant's set of responses to each of the Likert-scale affect questions was converted to Z-scores based on the mean and standard deviation of the participants' responses to that question. A Z-score is calculated for any score by subtracting from it the mean for the set of scores from which it is drawn, then dividing the result by the standard deviation of the set. Calculating Z-scores is a standard ESM procedure that allows comparing responses between students who use the Likert-scale differently. Essentially each person's score for a question is scaled in relation to their own average response for that question. Exploratory factor analysis on the responses confirmed that individual items grouped into four factors corresponding to the four constructs the survey was designed to measure: activation, self-efficacy, stress, and motivation. These categories of affect were calculated by averaging the Z-scores for each item within the construct.

Results

Differences in mean Z-scores were tested with two multivariate analysis of variance tests (MANOVA). The first test used two independent variables (non-school and school), and four dependent variables (activation, self-efficacy, motivation, and stress). It showed a statistically significant effect of activity type $F(4,1440) = 334.0$, $p < 0.001$. Subsequent one-way ANOVA tests contrasted activity type for each affect variable. They showed that there were statistically significant differences in mean Z-scores for school and non-school activities for all four variables, activation $F(1,1440) = 13.1$, $p < 0.001$, self-efficacy $F(1,1440) = 387.8$, $p < 0.001$, stress $F(1,1440) = 130.2$, $p < 0.001$ and intrinsic motivation $F(1,1440) = 1019.4$, $p < 0.001$. The second MANOVA used the same dependent variables but separated the school independent variable into non-STEM courses, STEM courses and IE physics. Non-school activities comprised the fourth independent variable. The MANOVA showed a statistically significant effect of activity type $F(12,1440) = 75.5$, $p < 0.001$. Subsequent one-way ANOVAs comparing the four activity types for each affect variable showed that all four affect variables had at least one statistically significant difference in mean Z-scores, activation $F(3,1440) = 9.15$, $p < 0.001$, self-efficacy $F(3,1440) = 140.5$, $p < 0.001$, stress $F(3,1440) = 57.8$, $p < 0.001$ and intrinsic motivation $F(3,1440) = 341.2$, $p < 0.001$. Post-hoc analysis was conducted using Tukey's HSD.

Results of the ESM measurements are summarized in Figure 2. Table 2 shows statistically significant differences and effect sizes in Z-score units. Both representations show that school and non-school activities were experienced very differently. School, compared to non-school, produced slightly more activation, much

lower self-efficacy, much higher stress, and much lower intrinsic motivation (i.e., more extrinsic motivation). Within school, STEM courses were experienced with slightly more activation and slightly more self-efficacy than non-STEM, but these differences were not statistically significant. IE physics was experienced, as compared to other STEM courses, with similar activation, but lower self-efficacy and higher stress. Intrinsic motivation was not different across school categories.

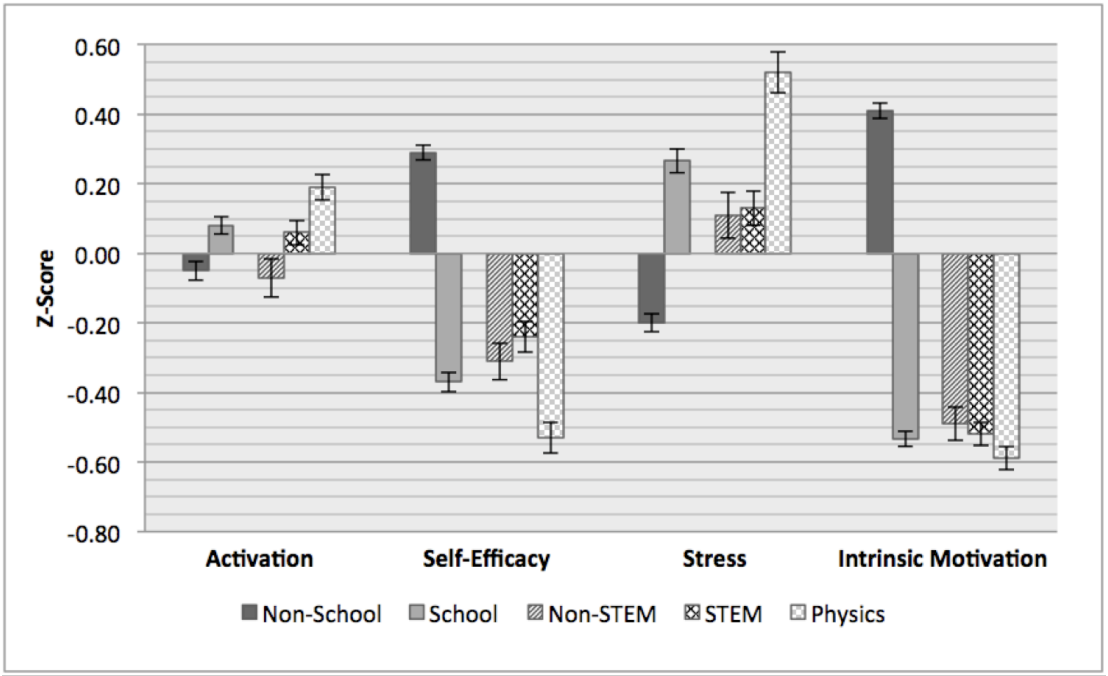


Figure 2. ESM data for all participants across the four activity types and four affective experiences. The vertical axis is in Z-score comparing the average experience in each activity to the overall average experience for each of the four affective experiences.

Table 2: Average experience, in Z-scores, across the four activity types with calculated differences between School V Non-School, Stem V Non-STEM, and STEM V IE physics. ***p<0.001.

	Non-School	School	School minus Non-School	Non-STEM	STEM	STEM minus Non-STEM	IE physics	IE physics minus STEM
N (surveys)	816	624		161	233		230	
Activation	-0.05	0.08	0.13***	-0.07	0.06	0.13	0.19	0.06
Self-Efficacy	0.29	-0.37	-0.66***	-0.31	-0.24	0.7	-0.53	-0.29***
Stress	-0.20	0.27	0.47***	0.11	0.13	0.02	0.52	0.39***
Intrinsic Motivation	0.41	-0.54	-0.95***	-0.49	-0.52	-0.03	-0.59	-0.1

Although the Z-score comparison is very useful and represents a common analytical approach to ESM, the fact that all comparisons are relative can be misleading. For example, the large negative Z-score for intrinsic motivation in school does not necessarily mean that intrinsic motivation was actually negative in school. Rather, students may have had a slightly positive intrinsic motivation in school that was nevertheless lower than a larger positive value for non-school experiences. For this reason, we examined the raw values for each of the affect categories to better interpret students' affective states within activity types. The following analysis is presented for motivation. A similar analysis was done for each of the other variables to ensure that the Z-score values in Figure 2 and Table 2 did not misrepresent students' affective states.

Intrinsic motivation was constructed from four questions on the ESM. Two of the questions, "Did you enjoy what you were doing?" and "How important was this activity to your future goals?" were measured on five-point unipolar scales. The other two questions, "How free or constrained did you feel?" and "How excited or bored did you feel?", were measured on seven-point bi-polar scales. The distributions of student responses on these four questions in school and non-school activities are shown in Figure 3. They show that intrinsic motivation in school was negative. School experiences tended to not be very enjoyable, and they were often

boring or unexciting. Feelings of freedom or constraint were more mixed but trended toward constraint. Finally, school activities were important to students' future goals. All of these results were in opposition to experiences outside of school, which tended to be enjoyable, free, and exciting while being less important to students' future goals. Based on these results, we concluded that the negative Z-score for students' experiences of intrinsic motivation in school activities was representative of motivation that was extrinsic, compared to intrinsic non-school motivation.

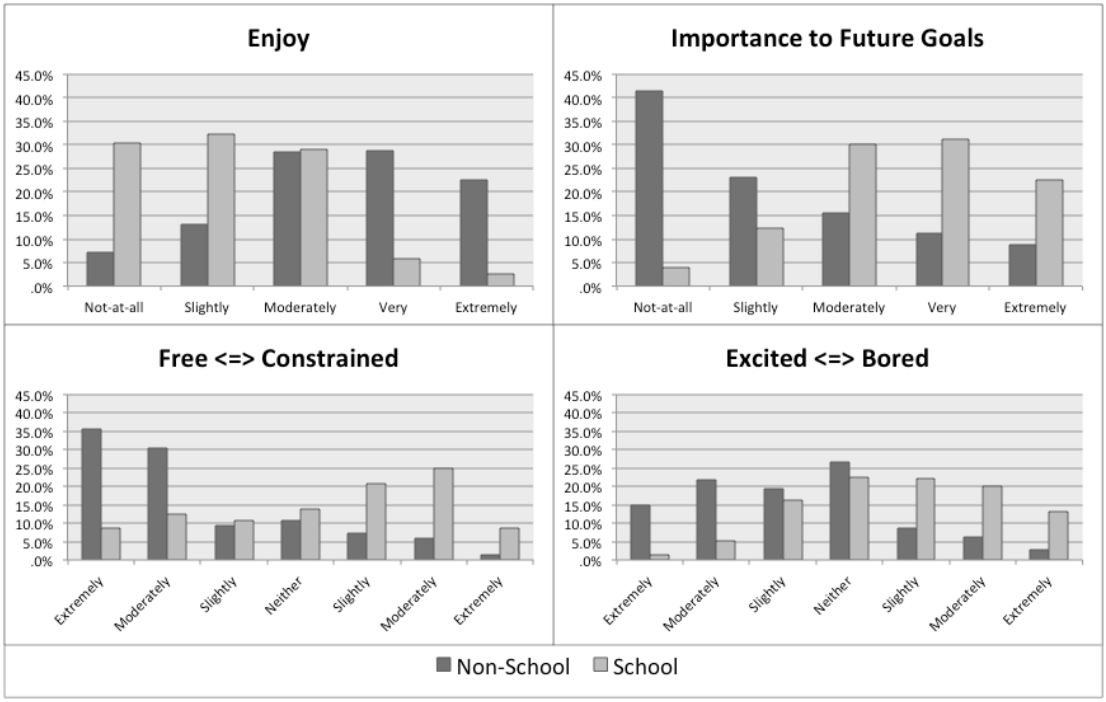


Figure 3: Distributions of student experience across the questions that made up intrinsic motivation.

Discussion

Comparing results for school and non-school, the differences in affect were consistent with the pressures that university students would be expected to feel as they strive to learn and generally perform well as novices in a competitive, (i.e., graded), environment. From this perspective, the lower self-efficacy and higher stress of school activities compared to non-school at least make sense. However, it is difficult to interpret the values of self-efficacy and stress for school activities. Perhaps these values could be improved; or perhaps they are a natural and healthy part of a student's struggle to learn. By contrast, the negative experience of intrinsic motivation in school activities (i.e., extrinsic motivation) is more clearly a cause for concern. Of course, extrinsic motivation is commensurate with the compulsory nature of many coursework activities such as studying for tests and completing problem sets. Educational systems generally tend to rely on structures and procedures, such as performance-contingent rewards, known to negatively impact intrinsic motivation (e.g., Deci, Koestner & Ryan, 1997). While the motivation results are not surprising, they are nevertheless worth paying attention to, since any educational program ultimately depends on students' self-regulated attempts to learn from what they are doing (Brophy, 2010). What the results of the present study show, compared to more theory-driven studies in the mold of Deci Deci, Koestner & Ryan (1997), is that negative intrinsic motivation is a measureable response to the education system, and not only a probable response based on the system's design. This "direct" measure of motivation has the potential to more squarely confront administrators, instructors, and other stakeholders with the consequences of motivation-reducing structures and procedures.

The IE physics course had higher stress and lower self-efficacy compared to other STEM courses, despite similar levels of activation (i.e., attentiveness, alertness, and concentration). These results could have been due to idiosyncrasies of the instructor. However the instructor's relatively long experience teaching the course, high average achievement, and distribution of grades similar to those of other STEM courses all undermine this possibility. We think it more probable that students' affect resulted from the experience of a fairly rigorous STEM course focused on learning conceptually challenging STEM content. If so, the present study would confirm the results of studies utilizing pre-post survey measures showing negative impacts of IE physics on self-efficacy (Kost-Smith, 2011). In the case of the present study, however, the negative impact of IE Physics was shown to occur within instructional activities, instead of upon reflection after the fact. This finding rules out the possibility that the negative impacts of self-efficacy measured by Kost-Smith may have occurred

primarily as a result of experiences outside instruction, such as when students received their grades. Rather, a negative effect on self-efficacy was located unequivocally within the process of instruction.

The lower self-efficacy for IE Physics might seem be inconsistent with the relatively high achievement in the course compared to other IE courses, since high achievement would be expected to encompass mastery experiences promoting self-efficacy (Bandura, 1997). However, while students in an IE course may learn more than they would in a traditional, lecture-based course (Hake, 1998), they are probably still far from mastery, which takes much longer to achieve than the brief experience of a semester-long course would allow (Chi, Feltovitch, & Glaser, 1981; diSessa, 2006; Ohlsson, 2009). Thus, it seems reasonable to assume that most students in a course would not be able to achieve anything near the levels of mastery that would lead to increasing in self-efficacy, even if they were learning at optimal levels. For this reason, it may be that educators should expect effective learning of challenging STEM content to be decoupled from self-efficacy on the time frame of one or two semesters during which most students experience this content. In essence, while students knowledge and skills are growing in the course, their awareness of what they don't yet understand and can't yet do are growing faster.

The present study has taken steps toward better understanding what is arguably a crucial and too long neglected aspect of STEM teaching and learning, affective experiences within instruction. The most important contribution of the study has been to measure the affective response in the midst of the instructional process. The results of this measurement, should they be replicated and extended in future studies, have the potential to raise the awareness of affective response to the process of instruction, and to motivate and inform the search for better instructional methods.

References

Adams, W., Perkins, K., Podolefsky, N., Dubson, M., Finkelstein, N., & Wieman, C. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics- PER*, 2(1).

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Belenky, D. & Nokes-Malach, T. (2012). Motivation and Transfer: The Role of Mastery Approach Goals in Preparation for Future Learning. *Journal of the Learning Sciences*, 21(3): 399–432,

Brophy, J. E. (2010). *Motivating students to learn*. New York: Routledge.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, *5*(2), 121–152.

Deci, E., Koestner, R., & Ryan, R. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. Psychological Bulletin, 125(6): 627–668.

diSessa, A. A. (2006). A history of conceptual change research. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 265-281). New York: Cambridge University Press.

Girod, M., Rau, C., and Schepige, A. (2003). Appreciating the beauty of science ideas: Teaching for aesthetic understanding. *Science Education*, 87(4): 574–587.

Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66 (1): 64–74.

Hektner, J., Schmidt, J., & Csikszentmihalyi, M. (2007). *Experience Sampling Method: Measuring the Quality of Everyday Life.* Thousand Oaks, CA: SAGE Publications.

Holleran, S., Whitehead, J., Schmader, T., & Mehl, M. (2010). Talking Shop and Shooting the Breeze: A Study of Workplace Conversation and Job Disengagement Among STEM Faculty. *Social Psychological and Personality Science*, 2(1): 65–71.

Koller, O., Baumert, J., & Schnabel, K. (2001). Does Interest Matter? The Relationship Between Academic Interest and Achievement in Mathematics. *Journal of Research in Mathematics Education*, 32(5): 448–470.

Kost-Smith, L. Characterizing, Modeling, and Addressing Gender Disparities in Introductory College Physics. PhD Thesis, 2011.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, 19(3): 355–365.

Lazarus, R. & Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Luzzo, D., Hasper, P., Albert, K., Bibby, M. & Martinelli, E. (1999). Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students. *Journal of Counseling Psychology*, 46(2): 233–243.

National Commission on Mathematics and Science Teaching for the 21st Century. (2000). Before It's Too Late. Washington: U.S. Department of Education.

Marra, R. & Bogue, B. (2009).Women Engineering Students and Self-Efficacy : A Multi-Year , Multi-Institution Study of Women Engineering Student Self-Efficacy. *Journal of Engineering Education,*

January.
Mujtaba, T. & Reiss, M. (2013). Inequality in Experiences of Physics Education: Secondary School Girls' and Boys' Perceptions of their Physics Education and Intentions to Continue with Physics After the Age of 16. *International Journal of Science Education*, 35(11): 1824–1845.
Ohlsson, S. (2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44*, 20–40.
Osborne, J., & Dillon, J. (2008). Science Education in Europe : Critical Reflections A Report to the Nuffield Foundation.
Sawtelle, V., Brewe, E., Goertzen, R. M., & Kramer, L. H. (2012). Identifying events that impact self-efficacy in physics learning. *Physical Review Special Topics - Physics Education Research*, 8(2).
Sawtelle, V., Brewe, E., Kramer, L., Singh, C., Sabella, M., & Rebello, S. (2010). Positive Impacts of Modeling Instruction on Self-Efficacy. PERC Proceedings, 289–292.
Semsar, K., Knight, J., Birol, G., & Smith, M. (2011) The Colorado Learning Attitudes about Science Survey (CLASS) for use in Biology. *CBE Life Sciences Education*, 10(3): 268–78.
Singh, K., Granville, M., & Dika, S. (2002). Mathematics and Science Achievement: Effects of Motivation, Interest, and Academic Engagement. *Journal of Educational Research*, 95(6): 323–332.
Stephens, N., Townsend, S., Rose, H., & Phillips, L. (2012). A cultural mismatch: Independent cultural norms produce greater increases in cortisol and more negative emotions among first-generation college students. *Journal of Experimental Social Psychology*, 48: 1389–1393.
Seymour, E. & Hewitt, N. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences,* Boulder, CO: Westview Press.
Thayer, R. E. (1996). *The origin of everyday moods: Managing energy, tension, and stress*. New York: Oxford University Press.
Thornton, R. & Sokoloff, D. (1998). Assessing student learning of Newton 's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula. *American Journal of Physics*, 66(4): 338–352.

Epistemic Networks for Epistemic Commitments

Simon Knight, Open University, sjgknight@gmail.com
Golnaz Arastoopour, University of Wisconsin-Madison, arastoop@gmail.com
David Williamson Shaffer, University of Wisconsin-Madison, david.williamson.shaffer@gmail.com
Simon Buckingham Shum, Open University, s.buckingham.shum@gmail.com
Karen Littleton, Open University, karen.littleton@open.ac.uk

Abstract: The ways in which people seek and process information are fundamentally epistemic in nature. Existing epistemic cognition research has tended towards characterizing this fundamental relationship as cognitive or belief-based in nature. This paper builds on recent calls for a shift towards activity-oriented perspectives on epistemic cognition and proposes a new theory of 'epistemic commitments'. An additional contribution of this paper comes from an analytic approach to this recast construct of epistemic commitments through the use of Epistemic Network Analysis (ENA) to explore connections between particular modes of epistemic commitment. Illustrative examples are drawn from existing research data on children's epistemic talk when engaged in collaborative information seeking tasks. A brief description of earlier analysis of this data is given alongside a newly conducted ENA to demonstrate the potential for such an approach.

Introduction

Seeking information in online environments is an increasingly important activity in a world in which students are no longer directed to pre-selected course books and materials. Yet, searching is a skill with which many – across age ranges – struggle. While support for technical aspects of searching may be of some assistance, we agree with Mason, Ariasi and Boldrin's claim that complaints regarding students' abilities at navigating the web are not technological, but rather epistemic issues around "the nature of knowledge and knowing, which may facilitate or constrain searching and evaluating sources of information on the internet" (Mason, Ariasi, & Boldrin, 2011, p. 139). This paper takes as its focus the seeking of information, claiming that this activity – particularly as mediated by and conducted through search engines – provides an *epistemic lens* through which researchers may explore the commitments learners make explicitly and implicitly about knowledge. These commitments are implicated in the ways students *select* sources, *use* them, and *make connections* between them in any information-seeking task. Analysis to explore issues at this level, and understand how to support learners to engage more effectively in their search practices is important. Both professional and academic learning contexts require high levels of information literacy; as such, an activity-oriented perspective on developing skills to support such literacy is an important contribution. This paper builds on recent contributions (including at ICLS 2012) calling for a shift from psychometric assessments for epistemic cognition, towards an exploration of the situated contexts in which epistemic practices are brought to bear. Specifically, we argue for a new approach: epistemic *commitments* – action-oriented ways of working – rather than *beliefs*, and the analysis of such commitments using Epistemic Network Analysis (ENA) to explore the connections between epistemic modes of information seeking. We suggest a focus on trace indicators of behaviors, and the connections between particular types of behavior (as analyzed using ENA) offer a productive new approach to the investigation of epistemic practices.

Seeking Information as an Epistemic Lens

In describing the established epistemic cognition literature, Mason (2009, p. 69) highlights broad agreement across models on the importance of two main facets – *what knowledge is, and how one comes to know*. Within the first area, two dimensions are noted: the *certainty* of knowledge (how stable or tentative knowledge is); and the *simplicity* of knowledge (how holistic a perspective of interrelated concepts, or simplistic a perspective of compartmentalized facts is taken). Similarly, Mason identifies two dimensions of the second area: the *source* of knowledge (from transmission to constructivism); and the *justification* for knowing (what warrants a knowledge claim – from authority to rules of inquiry). These models have informed analysis of the comprehension of multiple online sources – which may vary radically in the nature of their sourcing and justifications – in the understanding that students who regard knowledge as simple and finite may conduct brief and perfunctory searches with little recourse to integration or multiple sourcing (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). We thus agree that, "exploring students' thought processes during online searching allows examination of personal epistemology not as a decontextualized set of beliefs, but as an activated, situated aspect of cognition that influences the knowledge construction process" (Hofer, 2004, p. 43).

Research in this area indicates that students with more sophisticated inquiry stances are more likely to evaluate websites, and to do so meaningfully, while those with more sophisticated perspectives on the multiplicity of knowledge (or multiple perspectives) are more likely to integrate and critically evaluate multiple

online sources (Barzilai & Zohar, 2009; Bråten & Strømsø, 2006). While epistemic cognition is not a significant factor in understanding *converging* perspectives in online sources, for *conflicting* sources, those with evaluativist beliefs (who critique claims) perform significantly better in their comprehension (Barzilai & Eshet-Alkalai, 2013). A growing body of work associates search and sourcing patterns with particular patterns of epistemic metacognition (Mason et al., 2009), with think-aloud research indicating that students engaged in web-based learning spontaneously engage in epistemic reflection, particularly around source selection and credibility (Mason et al., 2011; Mason, Boldrin, & Ariasi, 2010), where students who verbalised about source credibility and information veracity significantly outperformed those who evaluated only sources (Mason et al., 2011). It should be noted, however, that the use of think-aloud protocols may increase such practices (Schraw & Impara, 2000; Schraw, 2000). We return later to the possibility that the *collaborative* search context may have higher external validity, as well as offering methodological advantages for gaining insight into a group's epistemic practices.

Situating Epistemic Commitments

The context of search is thus an interesting one for our investigations. Recent work has rejected an analysis of *beliefs* in favor of an action-oriented view:

> What we have called tacit epistemic beliefs might better be called epistemic commitments (Chinn & Brewer, 1993). Some theorists may be uncomfortable with the idea that one can have a tacit 'belief' that cannot be expressed, and the term epistemic commitment avoids reference to such beliefs. An epistemic commitment reflects a tendency to act in specified ways, such as a proclivity to provide justifications based on personal experience (Chinn, Buckland, & Samarapungavan, 2011, p. 146).

Furthermore, Sandoval (2012) has made related claims, calling for epistemic cognition researchers to take seriously a 'situated' approach:

> One important way to understand the epistemic ideas that people bring to bear is to examine their participation in practices of knowledge evaluation and construction. Changes in the form of participation are indicators of changes in the meaning that individuals make of the activity in which they are engaged. […] Change in participation can indicate a shift in epistemic perspective, but it is the shift itself that suggests what particular epistemic ideas are brought to bear in the first place (Sandoval, 2012, p. 350)

In a similar vein, Tsai (2004) suggests that information commitments involve "specific views about what counts as a successful explanation in the field (e.g. science) and […] general views about the character of valid knowledge or information" (Tsai, 2004, p. 105). Tsai (2004, p. 109) thus proposes three dimensions:

- *Standards for correctness:* evaluative standards ranging from 'authority' to 'multiple sources'
- *Standards for usefulness:* assessment of the usefulness of web-materials, ranging from 'functional' (e.g. ease of retrieval), to 'content' (e.g. relevance of retrieved information)
- *Searching strategy:* information-search strategy ranging from 'match' (of simple claims to questions) to 'elaboration and exploration'

While this turn from epistemic beliefs to commitments is an interesting one for analytics/data mining researchers who wish to analyze learners' behavioral traces as proxies for epistemic beliefs, it is still problematic, not least because as Wu and Tsai (2005) highlight, students may utilize both of the information commitments ('multiple sources' and 'authority') at the same time when evaluating the accuracy of the materials on the Web – a scenario which this framework does not have conceptual resource to explain. That is, while the orientations are proposed as dichotomous, or scalar, it is not clear that it is appropriate to think of them in such a way.

In our view, the action-oriented shifts described above are best characterized by the connections learners make between aspects of their sourcing behavior and information use. Thus, the focus should be on the emergence of information needs, and the use of multiple implicit and explicit criteria to assess the suitability of information for meeting those needs is dictated by a complex combination of searcher's action, task context, and technical mediation. Importantly, "…information seeking is not carried out for its own sake but to achieve an objective that lies beyond the practice of information seeking itself." (Sundin & Johannisson, 2005, p. 107). Therefore credibility assessments do not stand alone, but are connected to the continued seeking of information, and the ways in which information is used. Thus individual activities should not be considered in isolation: selecting multiple sources; claims around source authority; connecting pieces of information in complex ways; and so on, are not in themselves complex or simple. Context sensitivity is fundamental for a sophisticated epistemology; it is not very sophisticated to view the idea that the earth is round rather than flat as 'tentative' whereas theories of dinosaur extinction do require a more tentative stance (Barzilai & Zohar, 2012, p. 42).

Epistemic Frames

Epistemic Frame Theory (EFT) provides a means to conceptualize these connections between commitments. Epistemic Frames can be thought of in terms of the connections between elements usually described as: skills, knowledge, values, identities, and epistemological rules, from any particular domain. EFT is explicitly discourse oriented, and argues that an approach called Epistemic Network Analysis (ENA) may give insight into the frames of experts and novices working in a domain (Shaffer & Graesser, 2010; Shaffer et al., 2009). ENA thus offers a way to model the relations among elements of epistemic frames – which are constituted in discourse: particular facets of the frame (e.g. keywords indicating particular ways of working) become nodes, while connections between those nodes represent the patterns of connections between frame facets (e.g. the co-occurrence of keywords).

The search context is a particularly interesting one in which to deploy ENA. The theory takes as its unit of analysis any chunk of dialogue (a session) broken into meaningful chunks (stanzas). In the case provided in this paper we chunk stanzas by task, but for other analyses it may be more appropriate to chunk by search query. ENA allows us to model various types of connections, and broadly examine whether or not particular ways of making sense of information – in the confines of answering questions, or attempts at deeper understanding – co-occur with particular types of sourcing or connections between knowledge. Moreover, such analysis may offer insight into the quality of frame elements (nodes) – for example, claims about the 'authority' of websites might be rather trivial (for example, "it looks good") or more sophisticated (for example, "they used a scientific method"). Understanding how such justificatory elements of the frame are connected to sourcing elements may give insight into the pedigree of those sourcing decisions which would be missed by looking only for 'authority' claims. When we seek information we search for both in the sense that we search for information, and we search for a purpose; how users engage with those purposes is what matters, and how they connect those purposes to their epistemic commitments. Their sourcing decisions and the way they conceptualize the complexity of information is crucial. Thus, while search strategies matter, and an overreliance on individual (authoritative) websites or the consistent use of multiple websites (corroboration) might be of concern, their relationships to other epistemic assumptions are key.

In the work reported in this paper we take a previously analyzed dataset, and apply ENA to the epistemic discourse around searching for information to address a number of pre-assigned questions. We discuss the dataset further below, note though, that the use of this pre-existing dataset allows us to compare insights gained through close textual analysis, and those offered through ENA, thus supporting the development of a 'proof of concept' model for ENA around epistemic commitments.

The Collaborative Lens

A fundamental component of understanding the social context and role of language in learning is an analysis of how language mediates and represents learners' views on their learning. This component of learning is also fundamental to the theoretical and practical application of ENA, which takes as its data the discourse used in the course of students' learning practices. As noted above, it also avoids the methodological risk of artificially activating metacognitive strategies through the use of think-aloud techniques.

High quality collaboration also entails particular – epistemic – ways of working. In the context of epistemic commitments, take for example Hutchinson and Hammer's (2010) case study from a science classroom, in which framing by students which could be characterized as 'sensemaking' in nature (and, as we note below, accountable or exploratory) is contrasted with a more absolutist perspective. For example, at one point a student (Bekah) offers and explains an equation to illustrate her understanding – this is taken up and referred to collectively as "Bekah's Law", illustrating a cohesive tie (the repetition of terms through a text) demonstrating a type of common knowledge built up in that classroom (Hutchison & Hammer, 2010). This type of talk bears striking resemblance to exploratory or accountable talk, research on which focuses on the ways in which language is used "as a social mode of thinking – a tool for teaching-and-learning, constructing knowledge, creating joint understanding and tackling problems collaboratively" (Mercer, 2004, p. 137). In exploratory dialogue:

> Partners engage critically but constructively with each other's ideas. Statements and suggestions are offered for joint consideration. These may be challenged and counter-challenged, but challenges are justified and alternative hypotheses are offered. Partners all actively participate, and opinions are sought and considered before decisions are jointly made. Compared with the other two types, in exploratory talk knowledge is made more publicly accountable and reasoning is more visible in the talk. (Mercer & Littleton, 2007, p. 59)

In such talk, explanatory terms and phrases are more common, for example: *I think; because/'cos; if; for example;* and *also*. Similar characterizations of effective dialogue have emerged from the work of other researchers across a range of ages (Michaels, O'Connor, Hall, & Resnick, 2002; Resnick, 2001). This talk is thus explicitly epistemic, in that it embodies consideration of "the other's" perspective. The significance of this

type of dialogue for the study of epistemic commitments receives further support from Reznitskaya and Gregory (2013) who note that more sophisticated epistemic cognition of the 'evaluativist' variety, is closely associated with the kind of exploratory talk which is associated with educational gains. This claim – of an epistemic relationship to exploratory talk – is further supported in Rosenberg, Hammer and Phelan's work (2006). In that study, a case study was presented of a 15 minute discussion of the 'rock cycle' by a group of 8th graders – again, making use of dialogue excerpts to exemplify. Rosenberg et al., note that in the initial stages students were engaged in largely unproductive talk (there was some accretion of knowledge, with little explanation or evidence of understanding – it was largely *cumulative* in nature), suggesting this was because: "They [were] treating knowledge as comprised of isolated, simple pieces of information expressed with specific vocabulary and provided by authority" (Rosenberg et al., 2006, p. 270). After a brief intervention by the teacher, suggesting the students might build on their own knowledge, this talk instead shifts to more productive dialogue, seeking coherence and understanding in trying to create a theory and use terms they understand – the description, and excerpts provided here suggest this talk might be characterized as more 'exploratory' in nature. Exploratory dialogue is thus closely associated with a component of our approach to epistemic commitments around openness to ideas, and justification for them. This is particularly interesting given evidence that collaborative information seeking is a common activity (see Shah, 2012 for a review), and may have benefits for information seeking in classroom contexts (Lazonder, 2005) .

A Proposal for Epistemic Commitments

In addition to exploratory dialogue, other components of epistemic cognition are highlighted in the literature. Earlier we noted Mason et al.'s (2009, p. 69) claim that across models of epistemic cognition, there was a focus on the certainty, simplicity, source and justification for knowledge. We then noted Tsai's (2004, p. 109) framework for information commitments, comprised of: standards for correctness; standards for usefulness; and searching strategy. We thus recast these two positions such that our focus is on:

1. *Which sources of information are selected* – comprised of credibility decisions (from corroboration of information across sources, to trust in the authoritativeness of sources)
2. *How information is used* (in action – to justify claims, to make decisions) – comprised of justifications and source use (from dialogic approaches using talk of an exploratory nature, to attempts to directly approach questions by matching information to answers)
3. *How links between information are created (or not)* – comprised of claims, (explicitly in language and through structured environments, as well as implicitly through search patterns) made around connectedness of concepts (from a holistic to a piecemeal perspective of knowledge)

This recasting aligns well with the specific context being studied here – that of collaborative information seeking. It also provides three conceptually distinct (although probably empirically associated) constructs for study. In the two other models highlighted – Tsai's information commitments, and the general model described in Mason's analysis of the literature – it is not clear that each component can be conceptually distinguished. Specifically, 'certainty' in the general model seems likely to be a function of justification and simplicity. The sophistication of one's perspectives on 'certainty' depends on the purpose for which the information is being deployed, and the other information to which it is being associated – and indeed, whether one holds a complex enough view of knowledge to recognize the instability of certain information. Indeed, 'certainty' could be characterized as a connection between a facet of the information (publication date metadata, for example) and justification (recency, or information being well 'established' for example). Similarly, it is not clear that 'searching strategy' is a useful conceptualization of an information commitment given its strong relation to the tools at hand, and the type of task set and justificatory framework required for that task. The proposed model focuses on *whom we believe, how we justify claims; and how holistic a conception we have*. However, although 'epistemic commitments' recasts the constructs of other models, it still provides a lens for them. For example, 'certainty' is recast in light of our standards for credibility, explanation, and relating components of information such as new and old, or geographically located information; 'simplicity' is most clearly related to the third focus on connectivity; source to the first; and 'justification' to the second. Furthermore, the rhetorical shift both in the foci, and in the notion of 'commitments' over 'cognition' motivates an operationalization centered on:

1. *Source selection*, the corroboration of information across opened links, and the types of links repeatedly visited (e.g. use of authoritative sites such as 'BBC', repeated use of top links in search engine results pages, use of source metadata in the justificatory framework below).
2. *The type of justificatory framework used*, the assertion of information (perhaps closely related to a style of search which emphasizes precision of information with little consideration to its wider impact) versus reasoning and understanding activities (closely related to exploratory dialogue)

3. *The sorts of connections made* by students between concepts in their dialogue and document creation, and in the ways that users build links between information in their search patterns (building on search terms by rephrasing and appending new query terms, following internal links, and using terms from opened sources to find new ones all imply some commitment to holistic perspectives on knowledge).

This model thus describes both a conceptual and practical means to explore epistemic commitments in information seeking environments, and will be the model adopted in this work.

A Pilot Validation for Epistemic Commitments

The data described in this paper were taken from research the first author conducted in an English Secondary school with a pair and two trios of female 11 year old pupils (Knight & Mercer, Forthcoming). The researcher recorded an hour of discourse and (shared) display/browser use of the pupils while they were engaged in a set of assigned information seeking tasks around the topic of "role models". Some questions were closed ("How many women have won the Nobel Prize?") while others were more open ended ("Why do some people think Nelson Mandela is a good role model?"). In addition they were asked to justify their choices of information, and state their sources; these were questions explicitly designed to probe epistemic thinking. The data were analyzed for evidence of exploratory dialogue, and – making use of both the screencast and audio recordings – for epistemic behavior in pupil interactions with each other, and the information they sought. One of the groups was particularly unsuccessful in their performance (completing relatively few questions) while the other two performed rather better, although demonstrated some different ways of working. The original data were not 'coded' as such, rather a closer analysis using the methods of sociocultural discourse analysis (Mercer, 2004) was conducted. In the analysis here, this close focus on the properties of the text is used to motivate a shift towards coding utterances (turns by a single speaker) within stanzas (topically related sets of utterances; in this case utterances responding to set task questions) at the level of particular epistemic commitments by using key-terms (see **Error! Reference source not found.**).

Prior Analysis of the Data

As noted above, the data come from previous analysis with respect to epistemic cognition research, using a broadly sociocultural approach for a close hand-analysis of the transcript to explore the ways in which speakers make meaning together in context. In this section, we briefly summarize the findings of that research (which the first author conducted) before describing outcomes of ENA. There were marked differences in the behavior of the three groups. Group 1 in particular focused on how detailed sources were, and the repetition of keywords or information as indicators of usefulness, and had a general reliance on one website for many of the questions – although they talked very little about source quality, they spent some time discussing why their sources and information answered the questions and were 'useful'. In contrast, group 2 explicitly sought particular types of authority, noting the quality of BBC material, and potential problems with some sites (such as answers.com). Group 2 were very focused on extracting direct answers to questions from websites, and emphasized the novelty of information (i.e. "I didn't know that") as reasons for its importance often without directly addressing the part of the question asking them to justify their selection of information, or attempting to corroborate or make connections between bits of information. Group 3 showed the starkest difference in their behavior – and indeed, the poorest performance in terms of task completion. They emphasized *quantity* of information over *quality,* making no distinction between the qualities of different sources even where corroboration was attempted (e.g. treating 'answers.com' sites as equal authorities to the official website for the Nobel Prize). Preliminary analysis of this data by the first author in light of the theory of commitments proposed, highlighted salience of keywords for dimensions as in Table 1 which were used to code utterances for the application of ENA.

Table 1 – Dimensions of Epistemic Commitments for ENA

Code	Definition	Example
Source – Authority	Make reference to authorities when selecting information	"use BBC, that's a good site", "just use the top site", "it's a reliable site"
Source – Corroboration	Make reference to repetition of information when selecting information	"we said x cuz it was on lots of sites", "well everyone agreed"
Justification – Matching	Little attempt to sensemake in question; targeted matching of source information to questions.	"it's an answer", "look, it says it there", "what's it asking?"
Justification – Understanding	Attempt to make sense of information, use more accountable/exploratory dialogue key phrases (or epistemic superordinate codes)	"because", "so", "I think"
Simplicity – Simple	Make few connections between information, look for individual tokens of knowledge	"all the information", "need more facts", listing of claims without connections

Code	Definition	Example
Simplicity – Complex	Make connections between tokens of information, possibly across questions, contextualize information (e.g. account for temporal aspect), and make judgments regarding relative utility of information	"important information", making connections between claims (co-occurrence)
Search	Referring to search or webpage specific aspects of the task	"google that", "click there", "try searching for…"
General	References to general knowledge required (around role models in this case)	Keywords selected for general task relevance; e.g. "role models"
Specific	References to specific pieces of information in each question	Keywords selected for relevance to the specific questions asked e.g. "43 women"

Epistemic Network Analysis

Codes across epistemic commitments and search activities were applied at an utterance level (as examples in Table 1) with utterances grouped by question to form 'stanzas'. Co-occurrence of codes in stanzas are then identified to create an adjacency matrix – a quantification of connections between nodes. In the search study all three groups discussed at least the first four (of nine) questions (although one group did not complete it) thus only these four stanzas are analyzed. In ENA, connections are weighted by their presence across stanzas. As such, it is possible to place a threshold on connections such that only the most prominent connections are selected (and graphed when using ENA visualizations). In this case, the threshold is set to the highest level at which all three groups show any connections between nodes, a level largely dictated by group 3 whose utterances were characterized by simple phrases (many of which were off-task) and thus had very few connections between nodes. Furthermore, nodes which represent the greatest variation across groups are visualized on opposing axes. Thus by looking at connections and distance on axes, differences between groups can be explored.

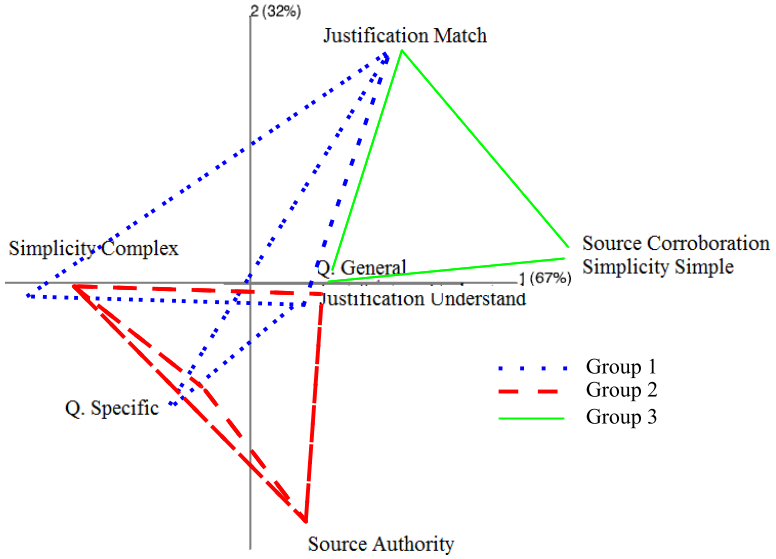


Figure 1. Epistemic Network Analysis for three groups' Epistemic Commitments

Figure 1 shows the ENA visualizations generated for Groups 1-3 (nodes have been combined to aid interpretation). Two dimensions accounted for most of the between-group variance. The x axis (dimension 1) accounts for most (67%) of the variance, with groups with low values making stronger connections with *Simplicity Complex,* and groups with high x values making stronger connections with *Simplicity Simple* and *Source Corroboration*. The y axis (dimension 2), then, largely distinguishes between *Justification Match*, and *Source Authority* with the use of question specific terminology. Here we see group 1 with larger y values representing stronger connections to matching, connecting to: understanding, making of *Question General* and *Specific* claims, and *Simplicity Complex*. Group one's talk might include things like: "we have to say *why* it matters, so [fact] is important because…". In contrast, we see group 2 with lower y values representing greater use of *Authority* – making the same connections otherwise (the cluster in the middle). This group's talk might be characterized by phrases such as: "the answer's [fact], because it's a good site".

The x axis defined the main difference between groups 1 and 2, and group 3 who made far fewer connections in general, and those that they did make were between: both *Corroboration* and *Simplicity Simple*, and making general claims, *Justification Understanding* and *Justification Match*. This might be interpreted as

indicating a perspective that simple knowledge obtained through corroboration is used to match general aims (and justified as such) although not actually targeting question specific knowledge claims (as is evident in the transcript and success rate with few targeted pieces of information associated with questions). Group 1 appears to be more focused on answering the questions using the information at hand, and making relationships between both the information found, and the questions. In contrast, group 2 appears to be focused on taking information from sources of authority that appear relevant to them. They still spend some time trying to relate this information and understand it, but they are not as focused on the requirements of the questions. It is interesting to note that there was little discussion of 'search' in any group, and that this is reflected in the lack of connections to this element of the frame. ENA thus offers a useful comparator for closer manual analysis. Through the analysis of connections in epistemic commitments, it gives many of the same insights as that closer analysis, while offering a method to scale analysis and provide real-time feedback.

Discussion

This paper set out to motivate a conception of epistemic commitments in the context of collaborative information seeking on the web, and to discuss the methodological and conceptual adequacy of ENA for their analysis. The example offered in this paper provides exemplifications of how such analysis might be conducted. Of course, ENA does not offer all the same insights as close manual analysis – for example, additional trace data would be required to identify specific websites used as authorities across contexts (e.g. 'BBC' here); further work will be needed to identify reliable key-terms. Some insights into other factors (such as novelty: "I didn't know that") are also lost, and some of these around unconnected commitments (as in the case of group 3's emphasis of quantity) should usefully be displayed in visualizations to indicate their presence as unconnected nodes. However, the brief descriptions offered from prior analysis of the data may be favorably compared with ENA results. Given the recoding for such purposes, this is arguably unsurprising. However, we would argue that we have avoided the risk of circularity; while more work is needed, this paper provides preliminary validation that ENA offers a *representational tool* for scalable interpretation of epistemic commitments, and that the notion of connections in epistemic frames is a productive characterization of epistemic commitments, offering more insight towards close qualitative [sociocultural] discourse analysis than simpler coding methods. However, we do not wish to overstate the suitability of this analysis in this case – much more work will be needed to define the interpretative space through which ENA for epistemic commitments is explored, including use of data designed for such use, validation and reliability assessment for interpretations of ENA output, and more work on providing text-oriented processing capabilities rather than the pre-selected cases provided here. However, given continued calls for the development of more situated, activity-oriented theories of epistemic cognition, and the specific aims of ENA for capturing the development of professional practices through enacting those practices, it may be well suited for analysis of epistemic activities such as information seeking. Comparison with 'expert' groups will also provide important comparison data; ostensibly while 'matching' is certainly an important connection insofar as it is core to actually addressing questions, it should not be a feature which defines the best quality of group activity (and indeed, groups 1 and 2 were very similar in many ways). Nonetheless, the conceptual scheme and preliminary analysis reported in this paper provide a development in thinking about analysis of epistemic commitments as *practice oriented* elements of working in the world. The next steps in this investigation which we are now planning are designed to assess a number of hypotheses, including:

1) In sourcing decisions, 'understanding' will give insight into the appropriateness of those decisions (corroboration v authority) that would not be gained through a simple analysis of either mode alone
2) Connections between task specific knowledge claims, sourcing decisions, and information seeking (search) will support analysis better than any of these alone, with respect to specific information claims
3) Connections between sourcing decisions and general knowledge claims (around broad task requirements) will give more insight than either alone
4) Connections between modes of the same dimension (for example, corroboration and authority) may be productively analyzed; positing decontextualized binary dimensions is problematic.

References

Barzilai, S., & Eshet-Alkalai, Y. (2013). The Role of Epistemic Thinking in Comprehension of Multiple Online Source Perspectives. In Y. Eshet-Alkalai, A. Caspi, S. Eden, N. Geri ,, Y. Kalman, & Y. Yair (Eds.), *Proceedings of the Chais conference on instructional technologies research 2013: Learning in the technological era*. Raanana: The Open University of Israel.

Barzilai, S., & Zohar, A. (2009). The Role of Epistemic Thinking in Online Learning. *Proceedings of the Chais conference on instructional technologies research 2009: Learning in the technological era*. Raanana: The Open University of Israel.

Barzilai, S., & Zohar, A. (2012). Epistemic Thinking in Action: Evaluating and Integrating Online Sources. *Cognition and Instruction*, *30*(1), 39–85.
Bråten, I., & Strømsø, H. I. (2006). Epistemological beliefs, interest, and gender as predictors of Internet-based learning activities. *Computers in Human Behavior*, *22*(6), 1027–1042.
Chinn, C., A., & Brewer, W. F. (1993). The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Review of Educational Research*, *63*(1), 1–49.
Chinn, C., A., Buckland, L., A., & Samarapungavan, A. (2011). Expanding the Dimensions of Epistemic Cognition: Arguments From Philosophy and Psychology. *Educational Psychologist*, *46*(3), 141–167.
Hofer, B. K. (2004). Epistemological Understanding as a Metacognitive Process: Thinking Aloud During Online Searching. *Educational Psychologist*, *39*(1), 43–55.
Hutchison, P., & Hammer, D. (2010). Attending to student epistemological framing in a science classroom. *Science Education*, *94*(3), 506–524.
Knight, S., & Mercer, N. (Forthcoming). The role of exploratory talk in classroom search engine tasks. *Technology, Pedagogy and Education*.
Lazonder, A. W. (2005). Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes. *British Journal of Educational Technology*, *36*(3), 465–475.
Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151.
Mason, L., Boldrin, A., & Ariasi, N. (2009). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and Learning*, *5*(1), 67–90.
Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the Web to learn about a controversial topic: are students epistemically active? *Instructional Science*, *38*(6), 607–633.
Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168.
Mercer, N., & Littleton, K. (2007). *Dialogue and the Development of Children's Thinking: A Sociocultural Approach* (New edition). Routledge.
Reznitskaya, A., & Gregory, M. (2013). Student Thought and Classroom Language: Examining the Mechanisms of Change in Dialogic Teaching. *Educational Psychologist*, *48*(2), 114–133.
Rosenberg, S., Hammer, D., & Phelan, J. (2006). Multiple Epistemological Coherences in an Eighth-Grade Discussion of the Rock Cycle. *Journal of the Learning Sciences*, *15*(2), 261–292.
Sandoval, W. (2012). Situating epistemological development. Presented at the International Conference of the Learning Sciences (ICLS), Sydney.
Schraw, G. (2000). Assessing metacognition: implications of the Buros symposium. *Issues in the measurement of metacognition*, 297–321.
Schraw, G., & Impara, J. C. (Eds.). (2000). *Issues in the measurement of metacognition*. Buros Institute of Mental Measurements and Department of Educational Psychology, University of Nebraska-Lincoln.
Shaffer, D. W., & Graesser, A. (2010). Using a quantitative model of participation in a community of practice to direct automated mentoring in an ill-formed domain. *Proceedings of the 4th International Workshop on Intelligent Tutoring Systems and Ill-Defined Domains held at the 10th International Conference on Intelligent Tutoring Systems (ITS '10)* (pp. 61–68).
Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., et al. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning.
Shah, C. (2012). *Information Retrieval: Collaborative Information Seeking: the Art and Science of Making the Whole Greater Than the Sum of All*. Springer.
Sundin, O., & Johannisson, J. (2005). The Instrumentality of Information Needs and Relevance. In F. Crestani & I. Ruthven (Eds.), *Context: Nature, Impact, and Role* (Vol. 3507, pp. 107–118). Berlin, Heidelberg: Springer Berlin Heidelberg.
Tsai, C.-C. (2004). Information commitments in Web-based learning environments. *Innovations in Education and Teaching International*, *41*(1), 105–112.
Wu, Y.-T., & Tsai, C.-C. (2005). Developing the Information Commitment Survey for assessing college and graduate students' evaluative standards for Web information and their searching strategies in Web-based learning environments. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (Vol. 2005, pp. 2500–2512).

Acknowledgements

The first author is grateful to the Epistemic Games group at UW-Madison and the Open University, who kindly supported the academic visit from which this work arose, and the participants and academic advisors for the initial work from which the data for this paper is drawn.

Framing Reflections on Instruction: A Precursor to Noticing

Vicky Pilitsis, Ravit Golan Duncan, Rutgers University, New Brunswick NJ
Email: pilitsisv2@yahoo.com, ravit.duncan@gse.rutgers.edu

Abstract: Noticing is the ability of teachers to attend and interpret student thinking to guide instructional design (van Es & Sherin, 2002). The skills involved in noticing can be challenging to develop in teacher education programs because of the cognitive load involved in attending to the context of the real classroom environment. Teacher education programs can thereby study a precursor to noticing, such as framing. Framing instruction involves developing a range of "seeing" events in the classroom. Thus, preservice teachers must frame their teaching experience in ways that privilege student thinking. In our investigation, we characterized the frames preservice teachers employed in their reflection paper. We found that preservice teachers who used frames that were more attentive to student ideas were more capable at analyzing student understanding in written student artifacts.

Introduction

Over the past three decades, standards documents have emphasized the importance in engaging students with the epistemology and practices of scientific inquiry (National Research Council [NRC], 1996, 2007, 2011). Since teachers mediate students' science learning, teachers must develop the knowledge and practices to implement inquiry teaching (Abd-El-Khalick & Lederman, 2000). An essential component in the implementation of inquiry-based teaching is the ability to attend to and interpret student ideas and to use such interpretations to guide instructional design (van Zee & Minstrell, 1997). The idea of attending to student thinking is not new- it has been a core aspect of pedagogical content knowledge (PCK) models for decades (e.g., Grossman, 1990; Magnussion, Krajcik, & Borko, 1999). More recently, Windschitl et al. (2012) advocated for instructional tools that support ambitious teaching, including practices that help teachers attend to student ideas. However, despite the need and benefits of attending to student thinking, this practice poses a major obstacle for experienced teachers and is even more difficult for preservice teachers (PTs) (Chamberlin, 2005). In particular, PTs struggle to make sense of student ideas and to develop these naïve ideas towards more normative understandings (Friedrichsen et al., 2009). Given these obstacles, we investigated whether and how PTs attended to student understanding in a science classroom.

Theoretical Framework

In every field, experts have the ability to notice and interpret events in their domain- they have "professional vision" (Goodwin, 1994). van Es and Sherin (2002) developed the framework of "noticing" to capture the notion of professional vision in teaching. The ability to notice consists of three sub-skills: (a) identifying what is important, (b) making connections between the specifics of classroom interactions, and (c) using what one knows about the context to reason about the classroom situation. Sherin et al. (2008) argued that the development of these skills poses challenges because classroom interactions, the fodder for noticing, are often fleeting and several occur simultaneously. It is even more challenging to help PTs develop these skills due to the cognitive load involved in attending to the messy contexts of real classroom interactions. Obtaining videos of PT instruction can also pose logistical challenges in some districts. To circumvent these issues, teacher educators can focus on the development of precursors to noticing. Framing instruction as being about student thinking may be a precursor to noticing. Hammer et al. (2005) termed frames as lenses to instruction and argued that framing involves developing a range of "seeing" events in the classroom. Therefore, a frame refers to expectations an individual has about a situation that affects what they notice and how they act. In order to develop the skills necessary for noticing, PTs must first frame their teaching experience in ways that privilege student thinking such that they observe these ideas and are subsequently able to interpret and respond to them. Levin, Hammer, and Coffey (2009) found that PTs have the ability to attend to student thinking, but what they notice in class depends in part on how they framed the lesson.

Most research that studied framing used video; however, in our investigation, we wanted to see if we could measure framing through written reflection papers. Researchers (i.e., Cavanaugh & Prescott, 2010; Hatton & Smith, 1995) have advocated for the use of reflection in teacher education programs as a vehicle to shift PTs attention away from themselves and towards attending to student thinking. Further, reflection fosters personal and professional growth, which has shown to improve teachers' knowledge and awareness of their classroom practices, including developing an awareness of student ideas (Baird et al., 1991). Engaging PTs in reflective practices could enhance the precursors of noticing, specifically framing, because how the reflection is framed places an emphasis on learning through questioning and investigation since it occurs after the lesson was taught,

thereby, eliminating all time constraints and other classroom pressures. By characterizing the frames that are expressed, teacher educators will be able to examine the various ways PTs attend to student ideas as well as study the development of framing, which could be a precursor to noticing. Therefore, our research questions for this study are:

- To what extent do PTs' framing of their lesson reflections account for student thinking?
- How do the frames expressed by PTs change over the course of a two-year teacher education program?

Methods

Study Context

This study was conducted in the context of a two-year Ed.M. certification program for secondary biology teachers. There were 16 PTs enrolled in the program. Four of the PTs were males and twelve were females. Fifteen of the PTs were Caucasian while one was of Asian descent. All of the PTs' undergraduate degrees were in the biological sciences with nine having biology degrees, three having animal science degrees, three having environmental science degrees, and one having a molecular biology degree.

The two year Ed.M. program included four life science methods courses that were taken in sequence (including a seminar that accompanied student teaching). The methods courses were geared to the development of knowledge and practices of model-based inquiry instruction. Each methods course had a slightly different focus. The first course, *Methods I*, focused on developing PTs knowledge of the nature of scientific inquiry. *Methods II* was a design course in which the PTs worked in groups to design an extended inquiry-based unit as well as implemented a short inquiry-based lesson. *Methods III*, which accompanied the student teaching internship, focused on the implementation of inquiry-based instruction as well as reflecting on their instructional methods. The majority of the PTs (15 out of the 16) completed their student teaching practicum in suburban high schools in the northeast while one of the PTs completed the requirement at an inner city high school in the northeast. Finally, the last course, *Methods IV,* engaged teachers in action research using data they had collected during their student teaching internship. The data that we used in this study was taken from *Methods II* and *Methods III.*

Data Collection

In this study, we used four assignments from *Methods II* and *Methods III*: (a) teaching experiment reflection paper from *Methods II*, (b) lesson set I and II reflection papers from *Methods III* and (c) reflective journals from *Methods II* and *Methods III.*

Teaching Experiment Reflection

The PTs were required to teach a lesson during the second methods course as part of their fieldwork and were asked to write a reflection paper about their experience. The reflection paper was divided into three sections. In the first section, the PTs were asked to provide a description of what went well and what did not go well in the lesson. In the second section of the reflection paper, the PTs were asked to select written student artifacts from the lesson and analyze the artifacts for student understanding in terms of scientific practices and content. In the third section of the paper, the PTs were asked to reflect on the revisions they would make to this lesson. In this study we analyzed the first two sections of the reflection paper.

Lesson Set Reflection I and II

During the third methods course, the PTs were asked to develop and implement two inquiry-based lesson sets during their student teaching practicum. Lesson set I was completed early in the semester (weeks 4-7) while lesson set II was completed towards the end (weeks 10-14). The lessons had to focus on model-based inquiry instruction. After implementing the lessons, the PTs were asked to provide a description of the lesson as well as to select written student artifacts from the lesson to analyze for student understanding in regard to scientific practices and content. We analyzed their descriptions on their lessons as well as their analysis of student artifacts.

Reflective Journals

The PTs were required to maintain a reflective journal throughout the two courses and to provide entries of about 250-300 words weekly. There were two types of journal entries: (a) answers to prompted questions that we asked several times during the course (i.e. what are the features of a scientific argument) and (b) personal and 'free-style" reflections on that week's class.

Data Analysis

We initially blinded all data sources in terms of PT and reflection paper. Using a constant comparative method (Glaser, 1965), we read through the sections of the reflection papers in which the PTs were asked to describe the previously implemented lesson. We noted any emergent frames (lenses) the PTs expressed. A frame was defined as the interpretative viewpoints PTs expressed while reflecting on lessons. For example, a focus on student participation or students staying on task would be categorized as an *engagement* frame. We identified six distinct frames, which we describe in the results section.

We then un-blinded the data to look for any trajectories of change in frames the PTs employed over the course of the teacher education program. We were interested in examining whether there were any clear patterns or shifts in the frames the PTs expressed. We noticed that three of the PTs did not hand in one of their reflection papers and another two of the PTs did not follow a clear pattern of change- they regressed and then progressed. Therefore, we selected the remaining eleven PTs for a more in depth analysis of shifts, seven of these PTs continually progressed towards framing instruction in ways that were more attentive to student thinking while the other four PTs selected did not progress (i.e., they either regressed continuously or used the same frame) in the frames they employed. For those selected PTs, we then analyzed the section of their reflection papers in which they were asked to analyze student understanding in written student artifacts that they had collected. We were interested in exploring whether those who employed frames that were more attentive to student thinking were more capable at identifying what students understood in the lesson. We read through that section of the reflection paper and using a constant comparative method (Glaser, 1965) noted any observed differences.

For the final part of our analysis, we wanted to examine whether the content of the different frames changed throughout the course of the teacher education program. We constructed tables for each frame according to the reflection papers (i.e., three tables for each frame) and highlighted the content that the PTs wrote about. For example, many PTs wrote about student participation when employing an *engagement* frame for all three reflection papers. The tables that we constructed consisted of the aspects of participation the PTs wrote about such as working collaboratively, quietness of students, and attentiveness of students in the teaching experiment paper and how students asking questions turned the lesson into a heated debate in lesson set II reflection papers. We reported the observed differences in the results section.

We triangulated the data by reading through journal entries that were written at the time of the implemented lessons (i.e., same time point as the reflection papers) to determine whether PTs expressed similar frames in the journals as were expressed in the reflection papers. We established inter-coder reliability by having two independent coders code the reflection papers (reliability ranged between 95-97%); any disagreements were resolved and codes were adjusted to reflect the consensus.

Results and Discussion

To What Extent to PTs' Framing of Their Lessons Account for Student Thinking

We identified six emergent frames described in table 1. We found that the PTs accounted for student understanding and student ideas to varying degrees. Table 1 presents the frames from the least to most attentive to student understanding. For example, PTs who employed the *engagement* frame focused on the students' interest and participation in a lesson with not much emphasis on student thinking. On the other hand, frames such as *scientific practices- students* and *building ideas* accounted for student understanding in either pedagogical practices or content. These results are encouraging because it revealed that PTs are able to attend to student thinking and that attention to thinking is a salient aspect of teaching for them. However, some PTs tended to express frames that did not account for student ideas, such as *activity sequence* frame. Davis (2006) argued that when PTs attend to learners their reflection centers on students' interest and motivation rather than learning content, which we also observed here. In our triangulation, we identified the same type of frames in both data sets (i.e., reflection papers versus reflective journals). In general, we found that PTs tended to express one frame or at most two. It also seemed that the PTs tended to employ the same frame regardless of the assignment (i.e., reflection paper versus journal prompts that were written around the same time as the reflection papers).

Table 1: Descriptions of emergent frames observed

Frame	Description of the Frame	Example from Reflection Paper
Activity Sequence	Characterized by a focus on providing a narrative or description of the lesson with minimal	"The students worked with the person they were sitting next to and talked about what they think happened in the story. After a little bit of time, volunteers read aloud their answers." (Nina, Lesson Set II)

Frame	Description of the Frame	Example from Reflection Paper
Scientific Practices- Teacher	Characterized by a focus on the teachers actions as they related to scientific practices, such as modeling, argumentation, etc.	"I handed each student modeling worksheets. Being as modeling is not something my students are familiar with, I felt it was necessary to help get them started so I wrote down the first two steps in the sequence of a fever with arrows on the board." (Jake, Lesson Set I)
Engagement	Characterized by a focus on student interest, participation, and staying on task.	"The story part went well, both periods were quiet, listening, and for the most part seemed interested." (Molly, Teaching Experiment Paper)
Accuracy	Characterized by a focus on obtaining the right answer. Teacher interprets student understanding in a binary way, either as correct or incorrect.	"The students demonstrated a basic form of knowledge of the content but did not go into much detail at all. A few students came up with the idea that antibodies were in the body." (Bani, Lesson Set I)
Scientific Practices- Students	Characterized by a focus on students' actions while implementing scientific inquiry practices such as modeling and evidence-based argumentation.	"None of the students used the data for generating evidence for claims, like viruses have various proteins…they [the students] failed to connect (link) data to evidence when making individual models." (Patrick, Lesson Set I)
Building Ideas	Characterized by a focus on taking students' knowledge and building upon it. Teacher interprets student's current level of understanding and describes possible connections to other content or suggests material to facilitate desired connections.	"By evaluating the worksheets I was able to provide material to help them [the students] more fully understand the implications of their solutions on the system as a whole by providing examples of previous attempts and solutions or additional data about the factors they involved." (Rachel, Teaching Experiment Paper

How Do the Frames Expressed by PTs Change Over the Course of a Teacher Education Program

PTs tended to express different frames throughout the teacher education program (Table 2). It seemed that initially the PTs tended to focus on the interest and participation of the students and there was a small shift towards focusing on student thinking. In general, we found that PTs tended to express one frame or at most two. It also seemed that the PTs tended to employ the same frame regardless of the assignment (i.e., reflection paper versus journal prompts that were written around the same time as the reflection papers).

Table 2: Frames observed according to reflection papers

Frame	Teaching Experiment Paper	Lesson Set I	Lesson Set II
Activity Sequence	0	3	3
Scientific Practices- Teacher	1	3	0
Engagement	13	3	3
Accuracy	2	3	0
Scientific Practices- Students	1	3	6
Building Ideas	1	1	3

We then looked at how individual PTs shifted throughout the course of the teacher education program (Table 3). We noticed that many of the PTs used different frames than they had employed in their first reflection paper. For example, during *Methods II* Sean wrote, "The lesson flowed smoothly. The students were engaged and the transitions between activities really caught their attention." Sean expressed an *engagement* frame because his focus was on the alertness and time on task of his students. Conversely during *Methods III*, Sean's focus completely shifted. He stated "The best part of the lesson was when the students were working with their models. They were able to construct logical representations of the material we had just covered." Here, Sean expressed a *scientific practices- student* frame because his focus was now on how the students constructed models rather than their interest in the lesson. We observed this change in both the reflection papers and the reflective journal entries.

Similarly, Jackie in her teaching experiment paper wrote "I found the students were very willing to participate in a respectful manner, raising their hands before speaking or calling out when there were no other hands raised." Like Sean, Jackie expressed an *engagement* frame because her focus was on the students' participation in the lesson. However, Jackie's lesson set II reflection paper had a completely different focus. In this reflection paper, Jackie was concerned with what the students were saying and how that related to their overall learning process. For example, she stated:

> "I made a concept map on the board but the map was really made entirely by the students as I would not write anything on the board until they discussed the ideas and concluded it was important to include. The students were able to take their initial ideas and elaborate and build upon them until they fully expressed their understanding."

In this example, Jackie expressed a *building ideas* frame because her focus in the lesson was now about the students elaborating and connecting their ideas. Further, we observed this shift in both the reflection paper and reflective journal prompts. It seemed that as the PTs gained more experience working with students in a classroom, they began to shift their focus from students being on task to becoming more aware of what students were actually saying and how their ideas related to core concepts. Thus, more classroom experience supported a shift from a focus on themselves to a focus on students' thinking (Berliner et al., 1988).

Table 3: Frames expressed by PTs throughout the teacher education program

	Teaching Experiment Paper	**Lesson Set I**	**Lesson Set II**
Christine	Engagement	Activity Sequence	Engagement
Patrick	MISSING DATA	Scientific Practices- Students	Engagement
Jackie	Engagement	Scientific Practices- Teacher	Building Ideas
Nina	Engagement	Engagement	Activity Sequence
Jack	Engagement	Accuracy	Building Ideas
Catherine	Engagement	Engagement	Activity Sequence
Sean	Engagement	Accuracy	Scientific Practices- Students
Nora	Engagement	Scientific Practices- Teacher	Activity Sequence
Molly	Engagement	Scientific Practices- Students	Scientific Practices- Students
Nadia	Engagement/Scientific Practices- Students	Activity Sequence	Building Ideas
Ava	Engagement/Scientific Practices- Teacher	Scientific Practices- Students	Scientific Practices- Students
Bani	Engagement/Accuracy	Accuracy	Scientific Practices- Students
Jake	Engagement/Accuracy	Scientific Practices-	Scientific Practices-

		Teacher	Students
Clare	Engagement	Engagement	Scientific Practices- Students
Rachel	Building Ideas	Building Ideas	MISSING DATA
Anna	MISSING DATA	Activity Sequence	Engagement

*We assigned a "missing data" code if PTs did not submit an assignment

We next wanted to explore whether the content of the different frames changed throughout the course of the teacher education program. We found that PTs began to be more elaborate and detailed in their descriptions of students' ideas in certain frames, specifically *scientific practices- students* and *building ideas.* Initially many of the PTs, who employed these frames, commented that the students had a difficult time explaining their models. They made statements like "the students did not explain or justify their models." However, in later reflections the PTs became more nuanced and explicit about *the ways in which* the students had difficulty using data stating that, "the students' content knowledge ability impacts how they understand the data and how they support their models" and "the students interpreted the data from the experiments and activities we performed in class and were able to incorporate this data to provide evidence based explanations." It seemed that the PTs began to see how different aspects of their students' learning process impacted their modeling skills. Overall, this shift was observed by all the PTs who employed *scientific practices- students* and *building ideas* frames in *Methods III.*

In addition, we found that there was a shift in the frequency with which PTs used students' responses in the form of quotes or comments from lesson activities when comparing the teaching experiment paper and the lesson sets. Initially, none of the PTs cited student responses in their reflection papers written for *Methods II,* while the majority of PTs used statements from students in their Lesson Set II reflection papers (Table 4). It seemed that as the PTs gained more experience in the classroom through their student teaching practicum, they became more aware of what students were saying and began to use the students' responses as evidence for justifying their reflections.

Table 4: Number of PTs who cited student responses in their reflection papers

	Teaching Experiment Paper	**Lesson Set I**	**Lesson Set II**
Number of PTs Who Used Student Responses	0	13	14

Although we did not observe a significant trajectory of change throughout the teacher education program, we wanted to explore what was different about the reflections written by the PTs who progressed towards framing instruction in ways that were more attentive to student thinking such as Sean, Ava, and Molly (to name a few). We found that in the earlier reflection papers, there were "seeds" of frames that were more attentive to student ideas. For example, in the teaching experiment paper, Sean's main focus was how the students participated in the lesson (*engagement* frame) but there were several statements about "bringing out student ideas." Additionally, Ava's main concern in the teaching experiment paper was about the students being on task during the activity (*engagement* frame) but Ava made several comments about "connecting ideas amongst topics as an effective instructional strategy". Overall, it seemed that in order to progress toward frames that were more attentive to student ideas there has to be an initial "seed" that becomes more prominent with experience.

For the final part of our analysis, we analyzed the section of the reflection papers in which the PTs were asked to analyze student understanding in the lesson. We wanted to explore if the PTs who progressed towards framing instruction in ways that were more attentive to student thinking such as Sean, Ava, and Molly (to name a few) were more capable of identifying what students understood or did not understand in their analysis as compared to PTs who did not progress in how they framed their lesson reflections such as Catherine, Nina, Nora, and Christine. In general, we found that PTs who progressed in how they framed instruction were more capable at identifying what students did not understand, commenting on students' prior knowledge or suggesting what topics should be stressed to learn the material. We also observed that the PTs were more interpretative of student understanding as the frames they employed were more focused on student thinking.

For example, in his teaching experiment paper, Sean, who progressed in the frames he employed, commented on what the students were not grasping in the lesson stating, "Every student verbally told me that onion cells do not have chloroplasts and elodea cells do, but many of their diagrams of onion cells included chloroplasts so I don't know where the disconnect is." Conversely, Catherine, who did not progress, commented

in the teaching experiment paper, "I think the students have a general understanding of what makes up a vertebrate" with no further elaboration. In both these instances, Sean and Catherine both employed an *engagement* frame when reflecting on their lessons but there is a significant difference in how they interpreted student thinking with Sean being much more specific about what the students were not understanding and provided evidence to support his claim.

In the final reflection paper, lesson set II, Sean, who employed a *scientific practices- students* frame, stated in his analysis of student understanding:

> "The models [that students drew] imply that the dots and lines among the required species are additional species (ancestors to humans), but the students do not label as such and do not explain them in their description. This shows that these students are missing a concept and are "filling in the blanks" of their understanding with these "dots and lines".

Here, Sean interpreted student understanding based on the lack of details in the students' models, which he believed indicated a missing connection amongst scientific ideas, in this case the evolution of species from common ancestors. In Catherine's final reflection paper (she employed an *activity sequence* frame), she stated, "Initially a majority of the students thought that bones are alive. They justified their opinion with correct ideas about the characteristics of living things." Catherine further elaborated her analysis by citing a student response from the lesson stating: "In Mala's model she wrote, "I think they [bones] are (alive) because when the body grows the bones grow too. All living things that have bones grow." However, although Catherine's analysis of student understanding was more evidence based (i.e., provided a response from a student) than in her initial reflection paper, she was still not as attentive to student understanding in regards to the students' overall learning process, whereas, Sean was much more analytical in his reflection by providing suggestions for what missing details in students' models suggested about their overall learning. In general, it seemed that the PTs who progressed in using frames that were more attentive to student thinking were more analytical in their examination of student understanding.

Conclusions and Implications

We observed a slight shift towards employing frames that were more attentive to student thinking as the teacher education program progressed. This finding suggested that teacher education programs should provide PTs with instructional tools when reflecting on lessons. Providing these tools will encourage PTs to employ frames that attend to student thinking and thus PTs will be more capable of analyzing student understanding in written student artifacts, an essential skill they will need in their future teaching career. Additionally, our findings suggested a methodological implication. In the past the notion of framing has been studied using videotape analyses but our results indicated that framing can be examined through analyses of written work, such as reflective practices.

These findings suggest that teacher education programs should provide PTs with instructional tools when reflecting on lessons. Providing these tools will encourage PTs to employ frames that attend to student thinking and thus PTs will be more capable of analyzing student understanding in written student artifacts, an essential skill they will need in their future teaching career.

References

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of the nature of science: A critical review of the literature. *International Journal of Science Education, 22*, 665-701.

Baird, J., Fensham, P., Gunstone, R., &White, R. (1991). The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching, 28,* 163-182.

Berliner, D. C., Stein, P., Sabers, D. S., Clarridge, P. B., Cushing, K. S., & Pinnegar, S. (1988). Implications of research on pedagogical expertise and experience in mathematics teaching. In D. A. Grouws & T. Cooney (Eds.), Perspectives on research on effective mathematics teaching (pp. 67–95). Reston, VA: National Council of Teachers of Mathematics.

Cavanagh, M., & Prescott, A. (2007). Professional experience in learning to teach secondary mathematics: Incorporating pre-service teachers into a community of practice. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice: Proceedings of the 30th Annual Conference of the Mathematics Education Research Group of Australasia*, Vol. 1 (pp. 182–191). Adelaide, SA: MERGA Inc.

Chamberlein, M. (2005). Teachers' discussions of students' thinking: Meeting the challenge of attending to students' thinking. *Journal of Mathematics Teacher Education,*8, 141-170.

Davis, E.A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Researcher,* 76 (4), 607-651.

Friedrichsen, P., Abell, S., Pareja, E., Brown, P., Lankford, D., & Volkmann, M. (2009). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching,* 46, 357-383.

Glaser, B. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12,* 436-445.

Goodwin, C. (1994). Professional vision. *American Anthropologist, 96,* 606-63.

Grossman, P.L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York/London: Teachers College Press.

Hammer, D., Elby, A., Scherr, R., & Redish, E. (2005). Resources, framing, and transfer. In J. Maestre (Ed.), *Transfer of learning: Research and Perspectives* (pp. 1-26). Greenwich, CT: Information Age Publishing.

Hatton, N., and Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education, 11,* 33-49.

Jacobs, V., Lamb, L., & Phillip, R. (2010). Professional noticing of children's mathematical thinking. *Journal of Research in Mathematics Education, 41,* 169-202.

Levin, D.M., Hammer, D., and Coffey, J.E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education , 60*, 2, 142-154.

Magnusson, S. Krajcik, J. & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95-133). Dordrecht, The Netherlands: Kluwer Academic Publishers.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Washington, DC: National Academy Press.

National Research Council (2007). *Taking science to school: Learning and teaching in grades K-8.* Washington, DC: National Academy Press.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academic Press.

Sherin, M., Russ, R., Sherin, B., & Colestock, A. (2008). Professional vision in action: An exploratory study. *Issues in Teacher Education, 17,* 27-46.

Simmons, P.E., Emory, A., Carter, T., Coker, T., Finnegan, B., Crockett, D. Richardson, L., Yager, R., Craven, J., Tillotson, J., Brunkhorst, H., Twiest, M., Hossain, K., Gallagher, J., Duggan-Haas, D., Parker, J., Cajas, F., Alshannag, O., McGlamery, S., Krockover, J., Adams, P., Spector, B., LaPorta, T., James, B., Rearden, K., & Labuda, K. (1999). Beginning teachers: Beliefs and classroom actions. *Journal of Research in Science Teaching,* 36, 930-954.

Tabachnick, B.R., & Zeichner, K.M. (1999). Idea and action: Action research and the development of conceptual change teaching of science. *Science Education,* 83, 309-322.

van Es, B., & Sherin, M. (2002). Learning to notice: scaffolding new teachers' interpretations of classroom interaction. *Journal of Technology and Teacher Education,* 10, 571-596.

van Zee, E., & Minstrell, J. (1997). Using questioning to guide student thinking. *Journal of the Learning Sciences,* 6 (2), 227-269.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education,* 96, 878-903.

"Teach Me How to Facebook!" Design-Based Research about Risk Prevention on Social Network Sites

Ellen Vanderhoven, Tammy Schellens, Martin Valcke
Ghent University, Department of Education, Henri Dunantlaan 2, BE9000 Ghent, Belgium
Email: Ellen.Vanderhoven@Ugent.be, Tammy.Schellens@Ugent.be, Martin.Valcke@Ugent.be

Abstract: Because of the increasing popularity of social network sites, the meaning of media literacy education has evolved. Indeed, in the 21st century, one of the important aspects of media literacy became to know how to behave safe online. However, most educational packages on the topic of e-safety have been developed without much theoretical consideration, and they have not been evaluated. Design-based research has been put forth as a good methodology to considerately develop effective educational materials. Therefore, in this study, a design-based research approach was used to develop educational materials about the risks on social network sites. By developing solutions based on existing knowledge and improving these solutions in 5 iterative cycles of implementation, evaluation and revision, this research results not only in effective practical solutions, but also in context-specific design guidelines that can be used by future researchers, practitioners and educational developers.

Introduction

With the rise of web 2.0, the meaning of media literacy has evolved. While traditionally, it referred to the ability to analyze and appreciate literature, the focus has been enlarged a few decades ago, thereby also including skills with regard to computers (Brown, 1998). Recently, this covers not only interactive exploration of the internet, but also the critical use of social media and social network sites (SNS) such as Facebook and Twitter. Since social media gives an excellent opportunity to create online content, the development of new skills is necessary. Livingstone (2004a) therefore describes media literacy in terms of four skills, this is as the ability to access, analyse, evaluate and create messages across a variety of contexts. It has been found that while children are good at accessing and finding things on the internet, they are not as good in avoiding some of the risks posed to them by the internet (Livingstone, 2004b).

In this respect media literacy education in schools has been put forth to empower teenagers (Livingstone & Haddon, 2009; Marwick, Murgia-Diaz, & Palfrey, 2010; Patchin & Hinduja, 2010). To encounter the increasing concerns about children's safety when using SNS, caused by for example privacy risks, sexual solicitation and cyberbullying, several prevention campaigns and awareness raising interventions have been developed (e.g., Insafe, 2014). However, most packages are developed without any theoretical base and a systematic review demonstrated that only few packages have been empirically evaluated (Mishna, Cook, Saini, Wu, & MacFadden, 2010). These scarce evaluation studies give evidence that a raise in internet safety knowledge is often achieved, but that evolutions in actual behaviour are much more difficult to obtain (Mishna et al., 2010). This is in line with the results of intervention studies in the more general field of media literacy education, that demonstrate that these interventions increase the knowledge about the specific topic of the course but have no impact on attitudes nor behaviour (Cantor & Wilson, 2003; Duran, Yousman, Walsh, & Longshore, 2008; Steinke et al., 2007).

As a reaction to the lack of theoretical base in interventions, the lack of theoretical implications of this intervention research, and the lack of evaluation studies in authentic settings, the design-based research methodology has been described (Phillips, McNaught, & Kennedy, 2012; The Design-based Research Collective, 2003). This methodology connects theoretical research with educational practice, and has been defined by Wang and Hannafin (2005) as

> a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practioners in real world settings, and leading to contextually-sensitive design principles and theories. (Wang & Hannafin, 2005, p. 6-7).

The output of design-based research contains both an increase of theoretical knowledge and a contribution to society, such as school programs (Reeves, 2006).

In the current research the design-based research approach lead us to develop effective educational materials to teach teenagers in secondary education how to use SNS safely (i.e., to increase awareness of risks and to change unsafe attitudes and behaviour) and to describe critical design guidelines for the development of

such materials. By developing solutions based on existing knowledge and improving these solutions in five iterative cycles of implementation, evaluation and revision, this research results not only in effective practical solutions, but also in a prototheory, describing context-specific design guidelines. The output of this research is therefore interesting for both researchers, developers and practioners (i.e., teachers).

The Initial Development of Solutions

As is typical for design-based research, design guidelines based on previous literature and theories were taken into account during the development of the initial materials. In this research, we took into account both general principles that are shown to be important in prevention campaigns (Nation et al., 2003) and more specific instructional design principles that follow out of the leading theory in education: constructivism (Duffy & Cunningham, 1996). This way, initial educational materials were developed for use in secondary education (Vanderhoven, Schellens, & Valcke, 2014a). The package consisted of a syllabus for the students and an instruction manual for the teacher. Every course could be carried out in one hour, thereby taking into account the request of teachers to minimize the duration of the lessons and the work load (Vanderhoven, Schellens, & Valcke, 2014b). The package aimed at both a raise in awareness about the contact risks on SNS, that is privacy risks, cyberbullying and sexual solicitation (DeMoor et al., 2008), and a decrease of unsafe behaviour on SNS after following the course. The different learning goals are described extensively in the teacher manual. All courses followed the same structure, as can be seen in table 1.

Table 1: The structure and content of the initial intervention

Structure of the course	Content
1. Introduction	The teacher introduces the topic, using the summary of risks (De Moor et al., 2008).
2. Two-by-two exercise	Pupils answer questions about a simulated SNS-profile on paper, together with a peer. These questions scaffold them towards different risks (cyberbullying, sexual solicitation and privacy risks).
3. Class discussion	Based on the answers of the pupils to the scaffolding questions and the answers given in the teacher manual, the teacher leads the discussion in class.
4. Voting cards	Pupils raise green or red cards, to show whether they agree or disagree with five given statements. Answers are discussed guided by the teacher.
5. Examples and theory	Some real-life examples are discussed. All the necessary information is summarized.

Method

The materials that were developed were put into practice in authentic secondary classrooms, while the impact on awareness, attitudes and behaviour of the pupils were measured using a pretest-posttest design. Based on the results of these measurements, the materials were refined. The revised materials were then implemented again in other classrooms. In total, five iterations of development, evaluation and refinement were conducted. The methodology was similar in the five intervention studies. However, some small changes have occurred. This is a typical characteristic of design-based research, where integrative research with varying methods is necessary to meet new needs and issues that emerge during the process (Wang & Hannafin, 2005).

Participants

The materials have been implemented in classes in secondary schools. In the first intervention study 1035 pupils participated, with a mean age of 15,14 (SD=1.88). In the second intervention study, 1487 pupils were involved, with a mean age of 14.9 (SD=1.11). In the third intervention study, 156 pupils followed a course, with a mean age of 15.39 (SD=0.61). The mean age of the 146 pupils that were involved in the fourth intervention is 12.92 (SD=0.61), and of the 80 pupils in the fifth intervention 15.64 (SD=1.23). In all studies, pupils were randomly divided over conditions, and no pupil participated in more than one study.

Procedure and Design

A pretest-posttest design was used in all intervention studies. This means that in all conditions, in all studies, pupils had to fill in an online pretest survey before the intervention took place. Afterwards, they followed the intervention, which was different in all studies. Finally, they filled in a posttest survey. In all intervention studies, a specific experimental intervention was compared with a control group. In the first two studies, no intervention took place in this control group, and pupils only had to fill in the surveys. In the last three studies, the intervention out of a previous phase was given to the control group, so that comparisons with the experimental group indicated the added value of the revised materials. The procedure is depicted in figure 1.

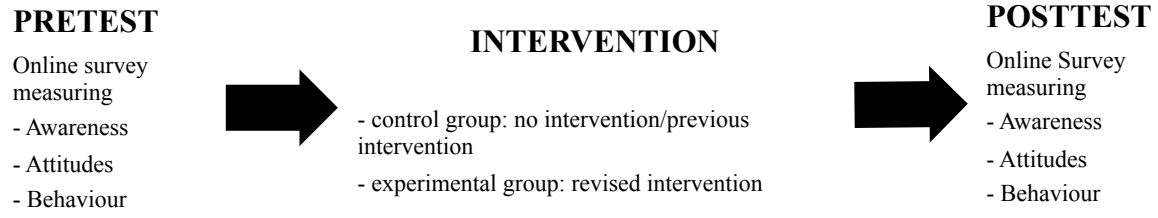


Figure 1. Procedure of the 5 evaluation studies.

Measures

Before and after the intervention, an online questionnaire was given to the pupils, measuring their awareness, attitudes and behaviour towards contact risks on SNS. This questionnaire was developed based on the contact risks as described by DeMoor et al. (2008). In the first two studies, three different scales were developed, one for awareness, one for attitudes and one for behaviour, all built on the base of the means of six or more items. They all had a satisfactory reliability as measured by Cronbach's alpha.

In the last three studies, the questionnaire was shortened, because pupils and teachers reported that it was too long and time consuming. Therefore, a new and shorter questionnaire was developed with less items on the awareness scale and with attitudes and behaviour measured based on the theory of planned behaviour following the manual of Fishbein & Ajzen (2009).

In all studies, an open question asked pupils about what they had learned during the intervention, to have a direct measure of increased awareness. Moreover, a direct binary measure of behavioural change was conducted by the question 'Did you change anything on your profile since the first questionnaire?'. If the latter was answered affirmatively, an open question about what they changed exactly gave us more qualitative insight in the type of behavioural change.

Analysis

Since our data clearly have a hierarchical structure, i.e. pupils in classes, the obtained data from pupils out of the same class might be dependent, and might so break the assumptions of simple regression analysis. In this respect Multilevel Modeling (MLM) is suggested as an alternative and adequate statistical approach. Consequently, since a significantly between-class variance could be observed indeed in the first two studies, a two-level structure is used: pupils (level 1) are nested within classes (level 2). The impact of the intervention on different posttest-scores -when controlling for the pretestscores- is evaluated by comparing the control-condition with the experimental conditions (i.e., adding the condition as a predictor in the model). Bonferroni corrections were used to control for multiple testing.

However, in the last three studies, no significant between-class variance could be observed, so there was no need to use MLM. Therefore, MANOVA's and multivariate repeated measures approaches have been used.

Results

Study 1: Implementation and Evaluation of the First Version of Materials

The initially developed materials were implemented and evaluated in authentic secondary classrooms. A positive impact of the given course on awareness could be observed, as pupils in the intervention group had an increased awareness about contact risks on SNS compared to the control group where pupils only filled in the pre- and post questionnaire. Yet, no impact of the courses on pupils' attitudes nor on their behaviour could be found analyzing the quantitative data. Still, if we analyze the qualitative data, some differences could be found. In the experimental group, significantly more pupils changed something on their profile than in the control group ($\chi^2(1)=15.60$, $p<.001$). As could be expected, pupils involved in the intervention about contact risks on SNS often changed their privacy-settings and their personal information such as their contact information. While these results indicate an impact of the course on the behaviour of a significant amount of teenagers, a lot of teenagers involved in the intervention reported not to have changed anything (i.e., 83%). Therefore, this first evaluation study demonstrated that the developed course had a significant impact on the awareness about the contact risks on SNS, but only limited impact on behaviour (Vanderhoven et al., 2014a).

First Revision of Materials

Following the results of the first study, several aspects of the intervention were inspected more closely. Based on observations made during the courses and comments of teachers about the intervention, possible holdbacks

were identified. For example, it was striking that during the class discussions and the voting game, well-liked pupils claimed that posting risky information "is not that bad at all". Following the theory of planned behaviour (Ajzen, 1991) -which states that behaviour is partly determined by the *social norm,* that is the social pressure people experience- and evidence that teenagers especially are sensitive to peer pressure (Sumter, Bokhorst, Steinberg, & Westenberg, 2009), it could be assumed that the well-liked peers described above had an important impact on other pupils' behaviour. Because of advantages of sharing information on SNS, such as communicating (Pruulmann-Vengerfeldt & Runnel, 2012) and creating an online identity (Hum et al., 2011; Madden & Smith, 2010), risky behaviour might be encouraged between peers and peer pressure might have prevented behavioural change.

Following this assumption, the educational materials have been modified. The possibilities for individual reflection were increased during the intervention by decreasing the 'peer time' that allowed pupils to be influenced by their classmates. This way, a larger impact on behaviour was aimed. Concretely, the two-by-two exercise was replaced by an individual task, forcing pupils to answer the questions about the simulated profile on their own. Subsequently, answers were discussed in class. The same revision was applied to the voting game. Pupils now had to reflect about the statements individually, instead of publicly raising green and red cards

Study 2: Implementation and Evaluation of the Revised Materials

Again, the materials were put into practice in an authentic classroom setting. In this study, there were two experimental conditions: in 43 classes the previous intervention was given (with collaborative learning), and in 25 classes the revised intervention was given (with individual reflection). Both groups were compared with the control condition where no intervention was given (43 classes). Both the original and the revised course had a positive impact on awareness as compared to the control condition ($\chi^2(1)=8.91$, $p<.02$ and $\chi^2(1)=7.24$, $p<.02$ respectively). However, only the revised course had an impact on attitudes and behaviour as compared to the control group ($\chi^2(1)=9.91$, $p<.02$ for attitudes, and $\chi^2(1)=5.67$, $p<.02$. for behaviour).

If we analyze the qualitative data, it could be found that while 7% of the control group changed something on their profile after the intervention, more pupils changed something in both the condition of individual reflection (13%, $\chi^2(1)=6.70$, $p<.01$) as in the condition of collaborative learning (17%, $\chi^2(1)=15.60$, $p<.001$). Pupils who changed something, often modified their privacy-settings, or adapted the personal information on their profile page.

The second evaluation study therefore showed that more time for individual reflection is valuable as a decrease of unsafe attitudes and behaviour could be observed compared to the control condition (Vanderhoven, Schellens, & Valcke, 2012).

Second Revision of Materials

After this second iteration, there was still room for improvement. Again, we took a closer look at the intervention and the remarks of teachers, pupils and observers. It was noted that the simulated profile in the course (the profile of "Sexy_Julie") contained so many risks that it was not experienced as realistic. Pupils might therefore feel like the risks are not relevant with regard to their own profiles. Therefore the materials have been revised. Where previously, students needed to complete an exercise with the simulated profile, now they had to make the same exercise with their own profile on a computer. This way, the authentic setting, that is described as an important instructional guideline out of constructivism (Snowman, McCown, & Biehler, 2008), was emphasized. Because computers were not always available at school, and because teachers reported that the course was already narrowly timed, the exercise was given as a homework task.

Study 3: Implementation and Evaluation of the Revised Materials

In this third study, there were two conditions. The control group (40 pupils) was involved in the intervention where the homework task was completed with the simulated SNS-profile of Sexy_Julie, while the experimental group (40 pupils) followed the course where the homework task was completed with the own profile. It was verified whether there was an added value of the revised materials concerning their impact on the awareness, attitudes, and behaviour with regard to contact risks on SNS. Three ANCOVA's have been performed with the posttest scores of respectively awareness, attitudes and behavior as dependent variables. There was no difference in impact between the two conditions when controlled for pretestscores, neither for awareness ($F(1,77)=.12$, $p=.73$), nor for attitudes ($F(1,72)=.001$, $p=.97$) or behavior ($F(1,72)=.38$, $p=.54$). On the contrary, qualitative data showed that the awareness of more different risks increased in the control condition where the profile of Sexy_Julie was used. Since there was no added value of using the own profile it was opted to keep the profile of Sexy_Julie in the package (Vanderhoven, Schellens, & Valcke, 2013).

Third Revision of Materials

For further improvement of the materials, we leaned back on the Theory of planned behaviour (Ajzen, 1991), because in our second study it has proven to be an interesting source of information to increase the impact of the intervention. Indeed, the social norm seemed to have an important impact on pupils' behaviour, as demonstrated by the larger impact of the intervention when the potential of peer pressure was decreased. Taking into consideration these results, it is notable that not only peers, but also parents have an important impact on the life of teenagers. Parents are often considered to be primary responsible for the moral socialization of the child (Maccoby, 2007) and play an important role in the education about online risks (Marwick et al., 2010; Pasquier et al., 2012; Safer Internet Programme, 2009). Moreover, Nation et al. (2003) emphasized that encouraging positive relationships, such as the relationship between parents and children, is a typical characteristic of effective prevention campaigns. Therefore, although peer pressure showed to have a negative influence on the impact of the intervention, parental involvement in school interventions might have a positive influence.

For this reason, the materials have been adapted, so that parental involvement is increased. Berkowitz and Bier (2005) described several strategies for schools to engage parents. One of the possibilities is to involve parents as clients, by organizing trainings in the topics of interest. Because of the rapid growth of SNS, many parents lack the expertise to guide and support their children's internet use (Livingstone & Bober, 2004). Therefore, training in internet related skills is necessary for parents as well. Following this assumption, parents were involved in the revised materials as clients by extending it with a parental information evening.

Study 4: Implementation and Evaluation of the Revised Materials

The revised intervention was again put into practice (Vanderhoven, Schellens, & Valcke, 2014c). However, since only 15% of the invited parents attended these parental information evenings, and only 19 of their children filled in both pre-and posttest, it was difficult to interpret results. Still, the qualitative results of this study show that the parental information session was useful to increase skills and literacy with parents. Moreover, most of the attending parents informed their children after the session about the risks on SNS and how to behave more safely. Therefore, it can be concluded that involving parents is effective at least to some extent. However, since only a limited amount of parents showed up at the information evening, organizing information evenings might not be sufficient to involve all parents. Although the attending parents were pleased with the information, and indicated that they learned a lot, there is no knowledge about the awareness, the internet literacy and skills of the non-attending parents. Analyzing the characteristics of the attending parents points to one of the main challenges of increasing parental involvement, this is involving all the parents and not only those parents who are already involved (Reynolds, 2005).

Fourth Revision of the Materials

Trying to encounter this challenge, new methods to involve parents were considered. Following Berkowitz & Bier (2005), the approach of actively involving parents as partners of the school was considered. This approach might be more appropriate to get in touch with all parents, and might so have a more positive influence on teenagers' behaviour. Therefore, the materials were revised, changing the homework task from an individual task into a task that needed to be completed in collaboration with the parents. Next to the questions that needed to be answered individually by both pupils and their parents, there were also a few questions that needed to be answered in communication, such as: on which questions did you have the same answer? Where did you/did you not agree? This way, all parents were actively involved in the intervention.

Study 5: Implementation and Evaluation of the Revised Materials

In this fifth study, an experimental group (9 classes) that followed the revised intervention was compared with a control group (7 classes) where the previous intervention was given. Both interventions had a significant impact on awareness about contact risks ($F(1,199) = 27.33$, $p < 0.001$), but especially boys benefited from the homework task with the parents when it came to a change in behaviour. While girls of both conditions posted less personal information and had less intentions to post personal or sexual information in the future, only the intervention with parental involvement showed this beneficial impact on boys.

Discussion and Conclusion

The last step of design-based research includes a reflection of the total research procedure and all findings, resulting in both practical solutions and improved theoretical understandings (Reeves, 2006). Therefore, we start this discussion with a conclusion and reflection on the findings. In a first phase of the research, educational materials have been developed based on different theoretical guidelines. Subsequently, important characteristics of effective educational materials were revealed in five iterative cycles of implementation, evaluation and revision. The final materials are therefore effective, as they increase awareness about risks on SNS, and decrease unsafe behaviour on SNS. Summarized it can be concluded that time for individual reflection and the involvement of parents in an intervention about risks on SNS is beneficial, especially for boys. However,

involving parents by organizing a parental information evening is not sufficient to engage all parents. Involving parents as partners, using a homework task, is put forward as a good alternative. Moreover, exercises with simulated profiles are just as good as real online profiles to obtain the objectives. Considering all these findings, a final effective and evidence-based practical solution was developed, that has an impact on both awareness and unsafe behaviour.

It is important that design-based research goes beyond designing and testing certain interventions. It must produce sharable *prototheories*, that help to communicate relevant implications to practitioners and educational developers (The Design-based Research Collective, 2003). Based on the results of this design-based research, context-sensitive design principles and theories are suggested. First of all, the initial design guidelines used to develop the materials need to be reconsidered. For example, in contradiction with the findings of Nation et al. (2003) that prevention campaigns need to be sufficiently dosed, it seems that an impact can be obtained already after a short-term intervention about risks on SNS. Second, collaborative learning, which was proposed as a central instructional strategy (Duffy & Cunningham, 1996), appears to be less successful in the case of reputation related behaviour like unsafe behaviour on SNS. Also, the value of an authentic learning context is put into perspective: while it seems to be valuable to include a SNS profile in the intervention, a simulated profile is sufficient to obtain results. There is no added value of making the context even more authentic, by including a real online SNS profile. Finally, the significance of positive relationships (Nation et al., 2003) is confirmed in the results of our studies: by including parents in the intervention, the impact on unsafe behaviour is enlarged. Combined with the finding that collaborative learning with peers is less effective because of the negative impact of well-liked peers, these results confirm the value of the impact of the social norm on behaviour, as stated by the theory of planned behaviour (Ajzen, 1991).

As a conclusion, it can be stated that this research resulted in both usable and evidence-based, educational materials about the risks on SNS and some contextually-sensitive design principles. The output of this research is therefore interesting for researchers, future developers and practioners. Moreover, this research is relevant to the theme of ICLS 2014 "Learning and becoming in practice" for several reasons. First of all, by focusing on changes in attitudes and behaviour, next to an increase in knowledge and awareness, we acknowledge that learning entails becoming a certain kind of person. We aim to influence pupils in a way that goes further than only increasing knowledge. By trying to make them critical citizens in this 21st digital century, and by teaching them to reflect about the different new risks that pop up with the rise of web 2.0, our materials aim to have a sustainable outcome. Second, by choosing for a design-based research approach, it was tried to gain insights by studying learning in real authentic settings, in this case the secondary classroom. By working in close collaboration with practitioners, materials have been developed that offer an opportunity for teenagers to learn, and become in practice.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. doi:10.1016/0749-5978(91)90020-T

Berkowitz, M. W., & Bier, M. C. (2005). Character Education Parents as Partners. *Educational Leadership*, *63*(1), 64–69.

Brown, J. (1998). Media literacy perspectives. *Journal of Communication*, *48*(1), 44–57. doi:10.1111/j.1460-2466.1998.tb02736.x

Cantor, J., & Wilson, B. J. (2003). Media and Violence: Intervention Strategies for Reducing Aggression. *Media Psychology*, *5*(4), 363–403. doi:10.1207/S1532785XMEP0504_03

De Moor, S., Dock, M., Gallez, S., Lenaerts, S., Scholler, C., & Vleugels, C. (2008). *Teens and ICT: Risks and Opportunities*. Belgium: TIRO. Retrieved from http://www.belspo.be/belspo/fedra/TA/synTA08_en.pdf

Duffy, T., & Cunningham, D. (1996). Constructivism: Implications for the design and delivery of instruction. In D. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 170–198). Simon and Schuster.

Duran, R. L., Yousman, B., Walsh, K. M., & Longshore, M. A. (2008). Holistic Media Education: An Assessment of the Effectiveness of a College Course in Media Literacy. *Communication Quarterly*, *56*(1), 49–68. doi:10.1080/01463370701839198

Fishbein, M., & Ajzen, I. (2009). *Predicting and Changing Behavior: The Reasoned Action Approach* (1st ed.). Psychology Press.

Hum, N. J., Chamberlin, P. E., Hambright, B. L., Portwood, A. C., Schat, A. C., & Bevan, J. L. (2011). A picture is worth a thousand words: A content analysis of Facebook profile photographs. *Computers in Human Behavior*, *27*(5), 1828–1833. doi:10.1016/j.chb.2011.04.003

Insafe. (2014). Educational resources for teachers. Retrieved from http://lreforschools.eun.org/web/guest/insafe

Livingstone, S. (2004a). Media Literacy and the Challenge of New Information and Communication Technologies. *The Communication Review*, *7*(1), 3–14. doi:10.1080/10714420490280152

Livingstone, S. (2004b). What is media literacy? *Intermedia*, *32*(3), 18–20.
Livingstone, S., & Bober, M. (2004). *UK children go online: surveying the experiences of young people and their parents*. London: LSE Research Online. Retrieved from http://eprints.lse.ac.uk/395/
Livingstone, S., & Haddon, L. (2009). *EU Kids Online: Final report* (EC Safer Internet Plus Programme Deliverable D6.5). London: EU Kids Online: LSE.
Maccoby, E. (2007). Historical overview of socialization research and theory. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization: theory and research* (pp. 13–41). New York: Guilford Press.
Madden, M., & Smith, A. (2010). *A Reputation management and Social Media.* Washington, DC: Pew Internet & American Life Project.
Marwick, A. E., Murgia-Diaz, D., & Palfrey, J. G. (2010). Youth, Privacy and Reputation (Literature Review). *Berkman Center Research Publication*, *5*, 10–29.
Mishna, F., Cook, C., Saini, M., Wu, M.-J., & MacFadden, R. (2010). Interventions to Prevent and Reduce Cyber Abuse of Youth: A Systematic Review. *Research on Social Work Practice*, *21*(1), 5–14. doi:10.1177/1049731509351988
Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention. Principles of effective prevention programs. *The American Psychologist*, *58*(6-7), 449–456. doi:10.1037/0003-066X.58.6-7.449
Pasquier, D., Simoes, J. A., & Kredens, E. (2012). Agents of mediation and sources of safety awareness: a comparative overview. In S. M. Livingstone (Ed.), *Children, risk and safety on the internet research and policy challenges in comparative perspective* (pp. 219–230). Bristol; Chicago: Policy Press.
Patchin, J. W., & Hinduja, S. (2010). Changes in adolescent online social networking behaviors from 2006 to 2009. *Comput. Hum. Behav.*, *26*(6), 1818–1821. doi:10.1016/j.chb.2010.07.009
Phillips, R., McNaught, C., & Kennedy, G. (2012). *Evaluating e-Learning: Guiding Research and Practice. Connecting with e-Learning*. Routledge, Taylor & Francis Group.
Pruulmann-Vengerfeldt, P., & Runnel, P. (2012). Online Opportunities. In S. Livingstone, L. Haddon, & A. Görzig (Eds.), *Children, risk and safety on the internet. Research and policy challenges in comparative perspective.* (Vols. 1-26, Vol. 6). Bristol, UK: The Policy Press.
Reeves, T. C. (2006). Design research from a technology perspective. In J. V. den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 52–66). London: Routledge.
Safer Internet Programme. (2009). *Assessment report on the status of online safety education in schools across Europe.* Retrieved from http://ec.europa.eu/information_society/activities/sip/docs/forum_oct_2009/assessment_report.pdf
Snowman, J., McCown, R., & Biehler, R. (2008). *Psychology Applied to Teaching* (12th ed.). Wadsworth Publishing.
Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2007). Assessing Media Influences on Middle School–Aged Children's Perceptions of Women in Science Using the Draw-A-Scientist Test (DAST). *Science Communication*, *29*(1), 35–64. doi:10.1177/1075547007306508
Sumter, S. R., Bokhorst, C. L., Steinberg, L., & Westenberg, P. M. (2009). The developmental pattern of resistance to peer influence in adolescence: will the teenager ever be able to resist? *Journal of Adolescence*, *32*(4), 1009–1021. doi:10.1016/j.adolescence.2008.08.010
The Design-based Research Collective, P. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*. Retrieved from http://www.aera.net/uploadedFiles/Journals_and_Publications/Journals/Educational_Researcher/3201/3201_DesignCollective.pdf
Vanderhoven, E., Schellens, T., & Valcke, M. (2012). Changing unsafe behaviour on social network sites: collaborative learning vs. individual reflection. *Manuscript Submitted for Publication.*
Vanderhoven, E., Schellens, T., & Valcke, M. (2013). How authentic should a learning context be? Using real and simulated profiles in an intervention about safety on social network sites. *Manuscript Submitted for Publication.*
Vanderhoven, E., Schellens, T., & Valcke, M. (2014a). Educating Teens about the Risks on Social Network Sites: Useful or Pointless? An intervention study in Secondary Education. *Manuscript Accepted for Publication in Comunicar*, (43).
Vanderhoven, E., Schellens, T., & Valcke, M. (2014b). Educational packages about the risks on social network sites: state of the art. *Procedia - Social and Behavioral Sciences*, *112*, 603–612. doi:10.1016/j.sbspro.2014.01.1207
Vanderhoven, E., Schellens, T., & Valcke, M. (2014c). Involving Parents in School Programs about Safety on Social Network Sites. *Procedia - Social and Behavioral Sciences*, *112*, 428–436. doi:10.1016/j.sbspro.2014.01.1185

Wang, F., & Hannafin, M. J. (2005). Design-Based Research and Technology-Enhanced Learning Environments. *Educational Technology Research & Development*, *53*(4), 5–23.

Acknowledgments

The research leading to these results has received funding from the Strategic Basic Research (SBO) Programme of the Flemish Agency for Innovation through Science and Technology (IWT). We would like to thank all the teachers involved in the study for their participation and dedication.

The Role of Identity Development within Tensions in Ownership of Science Learning

Jason Yip, Joan Ganz Cooney Center at Sesame Workshop, jyip@sesame.org
Tamara Clegg, June Ahn, Elizabeth Bonsignore, Michael Gubbels, Emily Rhodes, Becky Lewittes
University of Maryland – Human-Computer Interaction Lab
{tclegg, juneahn, ebonsign, mgubbels, charley}@umd.edu; emily@emilyrhodes.com

Abstract: Ownership of science learning is defined as learners being able to fully participate in the practicing culture of science, having greater control and possession over the ideas put forth, knowledge developed, and the science learning process. While ownership is beneficial to promoting science engagement, in this study, we show that conflicts in ownership of science learning manifest and can hinder learning. We document three focal learners who faced tensions and conflicts in their ownership of science learning. Specifically, we examine how learners' development and conceptions of ownership at home and school influenced how ownership of learning was expressed in an afterschool program called *Kitchen Chemistry* (KC). We argue that learners' expressions of ownership are a reflection of their identity development in science and that conflicts are a part of this manifestation.

Introduction

For many years, educational researchers have documented that many youth find aspects of traditional school science to be disengaging and irrelevant to their everyday lives (e.g., Atwater, 1996). Often in school science, teaching science is the equivalent of transferring knowledge from an authority (e.g., teacher, curriculum, software) to the students. Learners are often obligated to acquire knowledge from these credible and authoritative sources and later reproduce this abstract knowledge as correct answers (e.g., Fusco, 2001). Fusco (2001) argues that for science to be made relevant, learners need to engage in a *practicing culture of science learning* in which learners' own concerns, needs, issues, and experiences are brought to the forefront of learning. O'Neill and Barton (2005) contend that if learners were to have ownership of the science they were learning, they would be more motivated to engage. ***Ownership of science learning*** is defined as learners being able to fully participate in the practicing culture of science, having greater control and possession over the ideas put forth, knowledge developed, and the science learning process. Researchers argue that ownership can be a powerful way to support learners' engagement in science inquiry (e.g., O'Neill & Barton, 2005; O'Neill, 2010). The core assumptions of these studies are that having ownership in the learning process leads to greater motivation and participation in science learning.

However, within the literature there is an overly positive assumption that if science connected to learners' lives or encouraged active participation in a culture of science, this would help learners develop ownership of the knowledge and processes of science. Only a small number of studies have started to document the evolution of learners' ownership in science (e.g., O'Neill & Barton, 2005; O'Neill, 2010). Few studies have examined how social, personal, and cultural factors influence how a person interprets ownership and how an individual's own interpretation from one context (e.g., home) may change the dynamics of ownership in another context (e.g., school). In particular, as a learner transitions between different contexts, he or she may encounter conflicts in ownership of science learning, depending both social interactions and identity development (e.g., Pierce, Kostova, & Dirks, 2003). In this study, we document three focal learners who experience tensions and conflicts in ownership of science learning. Specifically, we examine how learners' development and conceptions of ownership at home and school influenced how ownership of learning was expressed in an afterschool program called *Kitchen Chemistry* (KC). We argue that learners' expressions of ownership are a reflection of their identity development in science and that conflicts are a part of this manifestation. Specifically, we ask two questions: 1) *what is the role of learners' identity development as they encounter tensions and conflicts in ownership of science learning?* and 2) *how do we best support learners' ownership in science learning, particularly in informal, project-based learning environments?*

Background

While there is consensus that learners' ownership can lead to higher engagement, researchers often examine ownership from different perspectives: individual and social. First, studies of ownership can take an *individual outcome* perspective; learners' ownership is a set of feelings and emotions that evokes a sense of control and possession within individuals and groups (e.g., Pierce et al., 2003). In their extensive review, Pierce and colleagues (2003) conceptually define *psychological ownership* as the "state where an individual feels as though the target of ownership or a piece of that target is theirs" (p. 5). Ownership is expressed in possessive emotions commonly associated with 'my', 'mine' and 'our.' Here, individuals might "feel" as though the target of

ownership is theirs. From an individual standpoint, learners' experiences, beliefs, goals, and cultural influences shape how ownership manifests. In this view, ownership is something that is achieved, and once achieved, ownership can provide a means for motivation and engagement in learning.

Ownership can also be examined from a *social process* standpoint. From this perspective, social contexts influence how ownership manifests; thus, ownership varies moment-to-moment for learners in various domains (e.g., Bandura, 2001). Ownership depends on the interacting relationships between learners, teachers, and the context. Although support of ownership needs to take place in the community, a lack of ownership can also be traced to power relations. Cornelius and Herrenkohl (2004) identify ownership of ideas as a manifestation of power in student-teacher relationships. Ownership of ideas implies a relation in power between individuals and concepts. In the realm of education, "whomever students perceive as having ownership of an idea - either themselves, their teachers, their textbooks, or their peers - will influence the relation that the student has to the idea itself" (Cornelius & Herrenkohl, 2004, p. 470). Therefore, students' expressions of ownership of knowledge are not standalone, but are tied up with the attitudes, participation and perspectives of adults and other learners. Using O'Neill and Barton's (2005) conception, our study acknowledges this duality: "Ownership is a dynamic and generative (social) process that exists in tension with ownership as an (individual) outcome" (p. 299). In this interpretation, there exists an acknowledgement that ownership is delicate and changing, but is still an innate part of people; thus, ownership exists as the dialectic between process and outcome and the dialectic between individual and social.

Ownership as Tensions and Conflicts in Science Learning

While most of the literature on ownership portrays the construct as an important and vital component to motivation and engagement (e.g., O'Neill, 2010; O'Neill & Barton, 2005), a limited number of studies in science education acknowledge the role of conflict in ownership of learning. Ownership of science learning often deals with power struggles between teachers and learners. O'Neill (2010) recognizes that classroom culture and structures impact learners' ability to take on ownership. Teachers and facilitators need to give up control and this is often a formidable challenge. Hay and Barab (2001) noted that learners' ownership of science learning conflicted with the actual practice of science. During their study, learners spent time working with actual scientists on an authentic investigation with real-life consequences. However, scientists needed to take control of the investigation away from learners to make sure the results were viable. Hay and Barab (2001) note the tension between ownership and authenticity that, "as authenticity increases ownership decreases" (p. 315). As the project became more authentic to a real-world community of scientists, the rules and cultural practices of science needed to be adhered to. Learners could not simply take control of the project and try out new practices.

Another aspect of control is placing structures, guidance, and scaffolds into science, which can diminish ownership. Reiser (2004) notes that in project-based science, great care is taken to contextualize the problem in learners' lives to support learners' ownership of the problem. However, project-based STEM learning is not full open inquiry. Similar to the authenticity issue, providing scaffolds and guides to those problems can take control away from the learners, thus weakening ownership. Finally, ownership of learning can cause learners to be overprotective of arguments and ideas. Haglund and Jeppsson's (2012) study directly examines the concept of ownership as pre-service science teachers learn to develop analogies for thermodynamics. As learners invested in their ideas and arguments, Haglund and Jeppsson found they became overly protective of the analogies.

Tensions in Ownership as a Reflection of Identity Development

As individuals become acculturated to the practices of science, they may encounter manifestations of power and social tensions that prevent them from taking full ownership. Within these documented conflicts of learner ownership in science learning, we do not yet fully understand the connection between social conflict and individual *identity development* and how to support learning in these moments. Identity development has a known strong connection to ownership. Organizational theorists Brown, Lawrence, and Robinson (2005) suggest that psychological ownership and self-identity are so correlated and tied to each other, that people mark and defend their territory as an extension of themselves. When individuals form strong feelings of ownership over physical or non-physical objects, they may attempt to mark these possessions exclusively as their own. If the possibility of infringement or threat to take ownership away from those objects occurs, individuals may engage in protective territorial behaviors that attempt to maintain levels of ownership. Pierce and colleagues (2003) call this threat, "the dark side of ownership" (p. 30). Both the processes involved in ownership and the innate characteristics of individuals leading to ownership are inextricably tied to how learners see themselves and are coming to see themselves. Ownership is thus tightly connected to learners' identity, but the connection has been rarely studied in the context of learning. In order to understand the role of ownership in learning, we need to not only recognize ways to promote ownership and identity development, but also remain aware of the limitations ownership can present to learning.

While studies have begun to tie ownership to identity development, few studies of ownership make direct connections to identity research. To better understand the role of identity development and conflicts in ownership of science learning, we utilize Wenger's (1998) conception of identity development through modes of belonging: imagination, engagement, and alignment. Modes of belonging are the ways participants see themselves as members of a community based on their engagement in practice, alignment in coordinated activities, and imagination of their world. *Engagement* is the process of how a member participates in the community. Through engagement, people work together to build relationships and communities of practice. *Alignment* is the process in which members take actions to align themselves to the goals and purpose of the community. Alignment is indicated through commitment, allegiance, and investment of energy. Alignment bridges space and time; participants can coordinate their energies, actions, and practice across other communities. Finally *imagination* is how members see themselves as connected (or not connected) to a broader community. Here, people imagine themselves as part of the community and gain a sense of connection with others. Imagination is broad, connecting to an extended identity. It involves seeing ourselves within a larger purpose and community. Using Wenger's identity framework, we shed light onto the conflicts arising in ownership and how we can best begin to address them to support learners' science ownership.

Methods

We employed the methods of a comparative case study (Yin, 2003) on a single implementation of Kitchen Chemistry (KC). KC is an afterschool program in which learners engage in scientific inquiry through cooking. We took on the role of participant observers; we both facilitated and observed the KC program. In the first four sessions of KC, learners engage in *semi-structured activities* to help prepare them to observe, reflect, and record food science activities. For this study, we specifically analyzed learners' participation on *Choice Days*. During these activities, learners are given opportunities to use what they have learned in KC to develop questions, hypotheses, and experimental procedures for their own food investigation. We observed learners making decisions on what recipes they want to modify, what variables they will control, what data to collect, and how to interpret their findings (Yip et al., 2012). Integrated into Choice Day was the use of several mobile apps. In this paper, we highlight *Scientific INQuiry* (SINQ), a social media app used by learners to develop and share questions, hypotheses, and investigation ideas (Gubbels, Yip, Kim, & Ahn, 2013). For this case study, we examined three focal learners and the conflicts that arose in KC based on their ownership of science learning. We analyzed learners' imagination, engagement, and alignment (Wenger, 1998) in science to investigate the role of identity development in ownership of science learning. We chose KC as a context for studying identity and ownership tensions because of the transformation in relationships of power (e.g., Cornelius & Herrenkohl, 2004); learners lead the investigation while facilitators played the supporting role.

Context and Data Collection

KC was implemented as a 12-week afterschool program that met once a week for roughly two hours in a local private school. Six learners between the ages of 8 to 11 participated in the program each week. The learners all attended the Montessori school that hosted KC. Each day we collected video recordings of all activities and discussions and software artifacts. We also conducted semi-structured interviews with four of the learners and their parents at two intervals of the program. In addition, we conducted interviews of the teachers of the focal learners and conducted classroom observations. Lead facilitators also recorded post-observational field notes of their experiences each day in KC. The facilitators in KC in the case studies are Beth, Emily, and Jason. We refer to the learners as Arman, Freddie, and Donna (pseudonyms).

Criteria for Case Selection and Data Analysis

We used the following three questions for the selection criteria: 1) What personal views did learners have of science?; 2) How did learners socially collaborate with each other and with facilitators?; and 3) What are the participation styles of the learners? Based on these questions, we selected three focal learners and vignettes that were representative of the present conflicts in ownership. We began the data analysis through an initial examination of the interview data, video recordings, software artifacts, and facilitator field notes. During this time, we wrote analytical memos and transcribed certain key portions of the data. Using methods outlined by Strauss and Corbin (2007), we used open coding to identify instances of social tensions, which included codes focused on learner distractions, social breakdowns, interruptions, arguments, difficulties in choice-making, frustration, and selfishness. We also coded for how learners engaged, aligned with, and imagined science. Using our analytical memos, photographs, and videos, we triangulated the data to determine if all pieces of evidence supported each other (Merriam, 2009). To establish validity in the coding scheme, we presented the codes to two external reviewers not closely involved with the study for an external code audit (Creswell, 1998). To make sure the cases were representative of ownership and conflict, we presented the case to the corresponding facilitator for validation. Finally, once the cases were thoroughly examined and developed, we conducted a cross-case analysis of the three cases to investigate similarities and differences in the data.

Key Findings

We begin each case with a description of the Choice Day activity that the learners and facilitators engaged in. In our analysis we then use Wenger's (1998) modes of belonging as a framework to analyze each case to understand the role of science identity development in ownership. Finally, we frame the conflict in each case through the individual and social processes perspective of ownership and its connection to identity development.

Case 1: Arman and the Spreadable Cookies

On Week 10 of KC, Arman, a 5th grade boy worked with Beth (facilitator) to create an investigation on spreadable cookies. Over the course of KC, we noticed that Arman tended to be quiet and did not always push for his own opinions. His teacher even expressed that Arman would often defer his choice and let others lead. In this Choice Day investigation, Arman was given the chance to follow through on an idea that he had initiated. Using SINQ, Arman entered his question: "What affects the spread of a chocolate cookie?" Arman wanted to pursue this cookie question and take control of the investigation. Beth also wanted to support his decision-making practices and cultivate his ownership over the spreadable cookies question.

Initially, Arman thought that butter affected the spread of the cookies. Beth suggested that they do a trial experiment to see which melts first, butter or vegetable shortening. Unfortunately, the pantry did not have any vegetable shortening to test out. Therefore, the duo needed to alter their plan. As they talked more about how to adjust the investigation, Arman showed Beth a website that might describe what the acid is in baking powder. She stated, "So without this, without the acid that is in this (baking powder), it should not rise." Arman pointed out, "So this should be flat (baking soda), flat cookies (points to baking soda) and not flat (baking powder)." Beth agreed and wrote down, "We should try one experiment with baking powder and one with baking soda and compare. We predict these will be flatter (baking soda) and we predict these will be fluffier (baking powder)." Beth called the final setup a "double or nothing" arrangement, in which their predicted more spreadable cookie consisted of liquid butter and baking soda, while their predicted less spreadable cookie was made with solid butter and baking powder. Although this was Arman's own investigation, he started to exhibit challenges with self-confidence. For instance, after this decision was made of the investigation setup, they needed to figure out the proportions for the recipe. Beth asked Arman, "Can you do the math and I'll type it in?" However, Arman looked hesitant and said, "You type it in, I'm not good at math." Being supportive, Beth said, "Oh well, we'll figure it out together." Even though the duo appeared to be making setup decisions together, Beth became concerned about her role as a facilitator. She stated she was worried the entire time about "taking over too much" or that he was not excited about the investigation. Beth claimed that, "*Arman might be opinionated, but you don't hear him voice his opinion*." Since it was difficult for Beth to interpret Arman's expressions of ownership, she acted cautious and did not want to overstep her bounds. Beth conveyed that she felt a tension in leading and supporting him.

Case 1: Analysis

In examining Arman's identity development, we must consider how his engagement and alignment tie together to his larger imagination of science and himself. With respect to engagement, Arman spent time working with Beth on the investigation. He was not distracted and his level of engagement with Beth did not decrease over time. He made investments into his cookie investigation. Arman aligned the cookie investigation to the practices of KC. The cookie investigation was not just about baking desserts. Arman spent time looking up ideas for his investigation in a cooking website and wanted to test out his question through an experimental design. However, even though Arman may have engaged and aligned with the practices of the investigation, he had difficulty taking complete charge of it. Specifically, we observed that Arman exhibited lower confidence; this might have affected how much control of the decisions he wanted, and how much ownership he exhibited. Wenger (1998) suggests that understanding imagination allows us to develop a more full picture of alignment and engagement. Towards the end of KC, we asked him if he could identify himself as a cook, designer, investigator, and/or scientist. He consistently reported what he called his "*slow progress*" with respect to these roles. Arman imagined people in these roles as being able to explain some knowledge or information to someone else, but expressed his limited imagination with respect to them in stating, "like I can't explain things really well." He reported he did not even think people at home and school would care to listen, "I just think like if I tell them and they don't really care, I don't know if they will really listen."

Arman's difficulty in imagining himself in these roles may have influenced his reluctance to take on stronger ownership of the investigations. As an individual aspect, Arman's outward behavior may have indicated a learner that took on aspects of ownership of the investigation, such having control over decisions in the food investigation and aligning his works through investments into the practices. However, in the social process of ownership, Arman also took on a deferential perspective to adults in his home and school life. This view of himself in comparison with adults might have made the choice-making process difficult for him. Although Beth attempted to support any decision he made to cultivate ownership, Arman may have wanted the "right" decision in KC or at the least decision he thought would make the adults happy.

Case 2: Freddie and the Greenies

On Week 06, Freddie, a 5th grade boy, was extremely excited to start his "Greenies" investigation. Since Week 03, Freddie had been clamoring to make green brownies (Greenies) as his Choice Day investigation. He worked with Emily as his facilitator. Upon the start of the session, Freddie immediately went to his station to begin. However, his enthusiasm soon deflated as he found out he had to fill out a goals chart, a scaffolded worksheet that learners fill out to determine what outcomes they wanted and what tasks they needed to accomplish their goals. Freddie, frustrated at even the notion of slowing down, raised his hands in the air to show his irritation, "Why don't I, I don't get this piece, cause it says what leavener should we use for taste and stuff? Like for texture? Seriously?" Emily, being patient, asked him, "What leavener should we use to make cakey brownies?", "Texture?", "Smell?" None of these were really pertinent questions to Freddie; all he wanted was to make brownies with a green color.

Instead, Emily wanted Freddie to consider how green food coloring would show up since all the other ingredients had different colors and tones. She suggested that they change something in the recipe to make sure the green food coloring shows up more. Freddie just wanted to add the green food coloring, "Let's just add it (green food coloring), just because it's (white chocolate) white, it will show up more." Emily again slowed him down, "Well hold on, that's the thing. Not everything in the recipe is white." Freddie argued back, "Brownies are brown because of the chocolate!" Needless to say, Freddie started to grow impatient with waiting and thinking. We observed that Freddie began to breathe heavily at this time; he wanted to go and just grab the ingredients. Emily asked him, "Are you getting frustrated?" to which Freddie nodded yes. She suggested they go get some fresh air. Emily also reminded Freddie they were a team and there was no rush to what they were doing. Once Freddie calmed himself a bit, Emily reminded him the Greenies might not be perfectly green and they need to consider how much white chocolate they would put in and how much green food coloring goes with it. Freddie stated that it does not matter how green it gets, "Anything green is good." Freddie was still excited, but was frustrated, "I just can't wait to start!"

Case 2: Analysis

While Freddie's frustration could be easily dismissed as impulsiveness, we argue that Freddie's identity development played a key role in his ownership and conflict. Freddie stated that he imagined that scientists and investigators 1) mix chemicals together in random ways; 2) serendipitously discover new substances; 3) make close observations; and 4) work in a lab that would be very similar to a kitchen. To Freddie, being a scientist meant doing a lot of hands-on mixing and making close observations of the final result. Wenger (1998) describes alignment as translating imagination into coordinated action. In KC, Freddie did not want to align with the slower and reflective practices emphasized in KC. Freddie expressed science activities at home (e.g., experimenting in the kitchen; determining if his cat lands on its feet) rarely had limitations, "At home you can choose, I can choose whatever I want" and "I like doing my own thing". In contrast, standing around and planning an investigation was not what he imagined scientists and investigators doing, and therefore, did not want to align to the KC practices. For engagement, Freddie wanted to take time to invest in the Greenies, but wanted to do this on his own terms. Freddie's goal was simple; he just wanted the brownies green. He wanted to start fast, get his hands into the cooking quickly, and not reflect on the investigation at hand. As part of his identity development, his reactions suggested he wanted to be known as the person that figured out how to make brownies green, not as the person that slowly planned the investigation.

In this case, science in KC and home came into conflict with Freddie's ownership, control, and imagination of roles. From the individual standpoint of ownership, Freddie wanted to keep his home science perspective of full autonomy and choice in the investigation. His perception of home science was of freedom of choice and supported his impulsive personality and bricoleur style of learning and engagement. Clegg and Kolodner (2007) describe Freddie's style of learning as a "bricoleur", one that investigates by manipulating objects and letting the product and learning emerge, often without planning. In contrast, from a social process standpoint, Emily took on a planner role; she preferred a more reflective and rule-based perspective. In this sense, when Emily wanted him to slow down and plan, she denied him full control and ownership over his investigation, even though she wanted to support his ownership. As Freddie's identity development took on the practices of how he conceptualized science learning, his ownership came into conflict with the more methodical and reflective inquiry practices of KC.

Case 3: Donna and the Puffy Cakes

Donna, a 5th grade female, wanted to make puffy cakes for her food investigation. She worked with Jason, a facilitator, for her Week 07 Choice Day. The investigation was focused on developing variations of cakes to determine what factors influenced cake density. As part of her ownership, Donna wanted to pursue the making of a cake. In SINQ, Donna entered the question "How do you make things (cakes) puffy?" and came up with the hypothesis that variations in eggs contribute to the puffy nature of a cake. Over the course of discussion, Jason suggested to Donna that they could do three egg preparations: eggs normal (yolk + egg white), egg white alone,

and the yolk alone. A fourth cake with just the batter alone (no eggs) would serve as the control since this was what the original recipe had indicated. Jason checked with Donna to see if this would work with her. She agreed with this setup. Meanwhile, Donna showed Jason that she could now crack an egg and separate the egg whites, a technique that she was very proud of.

Once the batter was mixed with the three different types of eggs in the cake ramekins, Jason set the mixtures into the oven. After the four cakes were baked, he brought them over to Donna. However, Donna started to become distracted by her friends. Jason tried to bring her back into the investigation by asking her to help him make measurements of the cakes. Donna noted that she was distracted by the noise and said she was tired of the measuring activity. In order to get her attention and reignite her ownership, Jason had her begin to taste some of the cakes. He had Donna try the control and she expressed enjoyment tasting it. However, as they continued with the measurements and observations, Donna appeared even more distracted. She looked tired and wanted to go play with an iPad™. Although others started to come and ask Donna questions about her cakes, Donna did not seem interested in answering them. Jason attempted to prompt her to wrap up her investigation, asking, "so which one is the most dense?" Again, she thought the egg white variation was the densest, but had difficulties articulating why. She thought the control and egg yolk versions were "too grainey" and that the egg white helped to enhance the flavor After all these observations were complete, Jason asked her, "So what do you think the egg is doing to all of these things right now that's different than the control?" Donna licked her fingers and shrugged her shoulders, indicating she did not have anything to say.

Case 3: Analysis

Similar to other KC learners, attention was a significant personal factor that influenced Donna's ownership of learning. Both Donna's mother and teacher brought up Donna's attention difficulties in interviews. As a result, Donna's ownership of the investigation waxed and waned as her attention shifted. However, attention alone does not provide the full story of what she chose to pursue in the investigation. We argue that understanding the roles of imagination, engagement, and alignment in identity development gives us stronger insight into her shifting ownership. First, Donna called herself a dreamer, someone that could come up with ideas and make them come true. For example, when we asked her about her career choices, cooking and designing careers always prevailed, but becoming a scientist always ended up last on her list of considerations. Like the bricoleur style of learning (Clegg & Kolodner, 2007) cooking and designing was meant to be free of restrictions. Donna's imagination of science was not due to lack of exposure to science; Donna's mother worked as a research scientist, studying allergies. During lab visits, her mother would often give Donna small experiment kits and let her play around with the equipment. She imagined that scientists constantly "make explosions" and "mix stuff" and they would inevitably "find cures and discover new things, stuff like that." Donna negotiated her engagement in science learning the way she imagined how she enacted the roles of scientist. She wanted hands-on mixing and exciting explosions, not the slower reflection and planning processes or the careful measurements and observations. Initially, when she was given the opportunity to come up with ideas, she took this task seriously. As such, Donna's alignment towards the hands-on aspects of sciences (e.g., mixing), the idea generation, and the end product development coordinated well between home and KC. However, when Jason and Donna began the slower reflection and measuring processes, she began to disengage.

Donna's conflicts of ownership over certain aspects of KC are a reflection of her identity through imagination, engagement, and alignment. From the perspective of the individual outcome, Donna chose targets of ownership that fit into her identity, such as science as hands-on activities that are constantly filled with "fizz", "explosions", and "fun". While Donna recognized that reflection and critical thinking were part of being a scientist, she did not think of herself as engaging in these characteristics. Instead, Donna wanted control and ownership over aspects she was familiar with from cooking at home and mixing in her mom's lab. Donna, who already had attention issues and imagined herself as a candy maker dreamer, had difficulties latching onto the slower paced tasks. From a social process standpoint, the conflicts in ownership are also a result of alignment between what Jason and Donna wanted. Wenger (1998) comments that since alignment concerns directing energy, it also concerns the power to exercise, inspire, and demand alignment. The coordination of actions constitutes shifts in identity and participation. However, in Donna's case, as her identity focused more on the fun scientist role, difficulties occurred as Jason tried support an alignment towards deeper reflection. As a result, Donna disengaged and disowned aspects of the investigation that did not fit with her identification.

Discussion: Connecting Identity Development to Ownership

In examining these three cases, we observed a variety of tensions in ownership of science learning. This study shows that while personal factors contribute to these conflicts, learners' existing identities and imaginations of science also play a significant role in how ownership is expressed. Each of these imaginations of science contributed to the tensions surrounding ownership. Arman imagined his identity development in science and design as slow going and not up to an imagined standard in which he could be confident to take control and ownership of his investigation, even with a supportive facilitator. In contrast, Freddie's case represents a learner

who wanted absolute control over his investigation. While impulsivity was a contributing factor to his ownership, for Freddie doing science meant having the ability to make any decision he wanted as long as it meant making his Greenies. Finally, in Donna's case, she faced challenges with attention, which illustrates her ownership as strong in the beginning, but waning in the end. However, she also imagined science as "fun" and full of explosions and color changes. When these dynamic occurrences did not happen, she chose to disown the investigation. Our work highlights the need to connect science learning to learners' existing imaginations about science. Learners like Freddie and Donna needed to see the fun parts of science – the explosions and the actions – to begin to align to the practices of the investigation. At points in the investigation, both Freddie and Donna needed to do science on their own terms for initial engagement.

However, supporting ownership of science learning is a complicated balancing act of authority and freedom. Even with their desires of engaging in the fun aspects of science, the learners needed some structure. Research has shown that reflection is powerful and necessary, both as learners are engaging (reflection in action) and later as they reflect on their engagement (reflection on action) (Schon, 1987). Typically schools designate the role of power to adults, while authority can be more shared in informal learning environments (O'Neill, 2010). This study reveals that issues of authority that appeared similar to formal learning were still an issue in our informal learning environment. Even though opportunities for learners to take more control were present, the facilitators still needed to think carefully about how to balance power and structure. We wanted learners to adhere to a culture of science, in which careful measurements and observations were paramount. However, some learners wanted to just cook and bake. Other learners wanted to do science the way they imagined it from their experiences at home, often in a very unplanned fashion. For these learners, placing guidance caused tensions in ownership. Some may advocate that more open inquiry environments in which learners have more control can promote ownership in learning. However, minimal guidance during instruction has been shown to be problematic due to cognitive load (e.g., Kirschner, Sweller, & Clark, 2006). Scaffolds and guidelines are still needed for learning, even in more open activities (e.g., Reiser, 2004). These cases reveal that simply giving learners a chance to control, possess, and own their activities is not enough for science learning. Without slower reflection and guidance, learners can miss important aspects of the science process.

Implications for Fostering and Supporting Science Ownership

Fostering learners' ownership can lead to deeper learning and engagement (e.g., O'Neill, 2010), but this study argues that cultivating ownership in science learning is complex due to the learners' identity formation in science. These cases show that one-size fits all approaches to supporting ownership did not exist since each learner's own identity development differed in trajectory. Some learners needed structured guidance to help make decisions and build confidence to take ownership, while others found guidance stifling. Our findings suggest that it is not enough to simply balance the amount of time and effort between fun activities and structured reflection to promote ownership. These two aspects need to be interconnected in ways that help learners see the necessity of both in science learning. For example, during interviews, Freddie spent time reflecting on the Greenies investigation and expressed the mistakes he made in the experimental setup. Freddie came up with new ideas about why his Greenies did not turn out well. He even worked with his mother on a more structured investigation at home to examine the differences between dark, milk, and white chocolates. Freddie needed an opportunity to cycle back and contemplate on the process and the importance of reflection.

Facilitators also need to be aware of the different needs of learners and be able to flow between structure and freedom dynamically. Quiet learners like Arman needed guidance from the facilitator to help develop ownership in his investigation. Arman expressed in the interviews he was overwhelmed with making choices in an inquiry environment. Without facilitator supports, it would have been even more difficult for him to take some aspect of ownership. For learners, like Freddie and Donna, who wanted to take stronger ownership over their investigations, we attempted to negotiate with them. We used strategies such as allowing them one set of ingredients they could "mess around" with, while another set would be used strictly for the investigation. Another strategy was to have learners switch roles from cook to technology recorder so that they could slow down and make closer observations. Supporting ownership was a give and take dynamic. Sometimes we needed to enforce structure so that learners could focus more. Other times, we allowed learners freedom to take risks and engage in experimental play in the ways they wanted. In all three cases, finding the right balance of control and ownership was delicate and did not always work the way the facilitators planned. Finally, this study shows that fostering ownership over the investigations was not always an immediate process and may not be long lasting. Arman needed time to gain confidence to take on more ownership of his investigation. Donna started strong in taking responsibility in the cooking aspects of the investigation, but her ownership dissolved quickly at the end, during the measuring and reflection process. For Freddie, it was only when he needed to figure out why his Greenies did not turn out well, he began to take on ownership of ideas and inquiry-based processes. We argue that fostering science ownership takes time and depends on how learners' identities and imaginations of science shift towards science dispositions.

Conclusion

This work shows that conflicts in ownership are not just an indication of authoritative power, but also a reflection of learners' identity development in science. We do not simply imply that facilitators surrender the responsibility of guiding the learning process to increase learners' ownership (e.g., Cornelius & Herrenkohl, 2004). Instead, we observed that each learner experienced science learning in diverse ways outside of KC. From these experiences, the learners developed different ways of imagination, engagement, and alignment in science, which ultimately influenced how they took ownership of their investigation and how they dealt with issues of power and social interactions. Our findings suggest that a delicate balance power between learners and facilitators is needed to foster engagement, imagination, and alignment, conducive to learners' ownership development. We make the argument that understanding the evolution of ownership of science learning also means further examining how learners' modes of engagement interact in multiple contexts.

References

Atwater, M. M. (1996). Social constructivism: Infusion into the multicultural science education research agenda. *Journal of Research in Science Teaching*, *33*(8), 821–837.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26.

Brown, G., Lawrence, T. B., & Robinson, S. L. (2005). Territoriality in organizations. *The Academy of Management Review*, *30*(3), 577–594.

Clegg, T. L., & Kolodner, J. L. (2007). Bricoleurs and planners engaging in scientific reasoning: A tale of two groups in one learning community. *Research and Practice in Technology Enhanced Learning*, *02*(03), 239–265.

Cornelius, L. L., & Herrenkohl, L. R. (2004). Power in the classroom: How the classroom environment shapes students' relationships with each other and with concepts. *Cognition and Instruction*, *22*(4), 467–498.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Fusco, D. (2001). Creating relevant science through urban planning and gardening. *Journal of Research in Science Teaching*, *38*(8), 860–877.

Gubbels, M., Yip, J., Kim, J., & Ahn, J. (2013). Scientific INQuiry (SINQ): Social media for everyday science learning. In *Proceedings of the iConference 2013* (pp. 1102 – 1105). Fort Worth, TX.

Haglund, J., & Jeppsson, F. (2012). Using self-generated analogies in teaching of thermodynamics. *Journal of Research in Science Teaching*, *49*(7), 898–921.

Hay, K. E., & Barab, S. A. (2001). Constructivism in practice: A comparison and contrast of apprenticeship and constructionist learning environments. *Journal of the Learning Sciences*, *10*(3), 281–322.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley and Sons.

O'Neill, T. (2010). Fostering spaces of student ownership in middle school science. *Equity & Excellence in Education*, *43*(1), 6–20.

O'Neill, T., & Barton, A. C. (2005). Uncovering student ownership in science learning: The making of a student created mini-documentary. *School Science and Mathematics*, *105*(6), 292–301.

Pierce, J. L., Kostova, T., & Dirks, K. T. (2003). The state of psychological ownership: Integrating and extending a century of research. *Review of General Psychology*, *7*(1), 84–107.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, *13*, 273–304.

Schon, D. A. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Strauss, A. L., & Corbin, J. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 3rd ed*. SAGE Publications.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge Univ Pr.

Yin, R. K. (2003). *Case study research: Design and methods* (Vol. 5). Thousand Oaks, CA: SAGE.

Yip, J. C., Clegg, T. L., Bonsignore, E., Gelderblom, H., Lewittes, B., Guha, M. L., & Druin, A. (2012). Kitchen Chemistry: Supporting learners' decisions in science. In J. van Aalst, K. Thompson, & P. Reimann (Eds.), *Tenth International Conference of the Learning Sciences* (Vol. 1, pp. 103–110). Mahwah, NJ, USA: Erlbaum.

Acknowledgments

We want to thank the CI Fellows program for funding this work. We would like also to thank the children, teachers, parents, and staff of The Green School for their participation.

Developing Mechanistic Model-Based Explanations of Phenomena: Case Studies of Two Fifth Grade Students' Epistemologies in Practice over Time

Christina V. Schwarz, Li Ke, May Lee, and Joshua Rosenberg
Department of Teacher Education, Michigan State University, 620 Farm Lane, East Lansing, MI 48824
cschwarz@msu.edu, keli1@msu.edu, leemay1@msu.edu, jrosen@msu.edu

Abstract: To foster meaningful engagement in scientific practices within classrooms, we must better understand how students can productively develop their epistemologies in practice (EIP) across contexts. This study traces how two high-performing fifth grade students engaged in scientific modeling across three modeling-centered units over one and a half years. To analyze their epistemologies in practice over time, we examined their model-based explanations and reflective talk about the rationale and purposes of their explanations. While both students developed more mechanistic explanations of phenomena, they did so in different ways. One developed a meta-level rhetorical strategy to explain "how and why" phenomena occurred, even in non-prompted contexts. The other used a reductionist analytical strategy to look for deeper level mechanisms of phenomena. These cases provide evidence that students develop EIPs across contexts, and leads to insights about what develops, what might influence this development, and how EIPs can be further supported in classrooms.

Supporting Meaningful Engagement in Scientific Practices

The science education community has taken a "practice turn" with respect to K-12 reform efforts (Ford & Forman, 2006). The Framework (National Research Council, 2012) and NGSS (National Research Council, 2013) have highlighted the idea that students should learn core disciplinary knowledge while engaged in scientific practices. This turn recognizes the importance of engaging students in communities of practice (e.g., Wenger, 1998) that pursue classroom versions of the knowledge-building goals embraced by the scientific community (e.g., establishing claims supported by empirical evidence, and generating, evaluating, and revising knowledge products that embody descriptions and mechanisms of phenomena). But for these goals and related practices to be meaningful (rather than procedural), they must be guided by the *epistemological considerations* that characterize disciplinary science (e.g., Duschl, 2008). However, engaging students in epistemologically meaningful practices in K-12 classrooms is extremely challenging. Furthermore, the science education and learning sciences communities do not yet have strong understandings of how students develop their epistemologies in practice over time (across units and classroom settings) and what increasingly productive epistemic engagement might look like. Epistemological considerations are critical for engaging in scientific practices in ways that are meaningful for students and authentic to the discipline. If we are to successfully enact reform efforts emphasizing student engagement in scientific practices, then we must better understand how students learn to use epistemological considerations to productively engage in practices over time.

Studying Students' Developing Epistemologies in Practice

Our work has centered on how students learn to engage in scientific practices within elementary and middle school contexts (Schwarz et al., 2009). In particular, we have been studying how elementary and middle school students engage in scientific modeling, explanation, and argumentation practices to make sense of the world and the role of epistemological considerations in making those practices meaningful. By epistemological considerations, we refer to students' practical epistemologies as they are engaged in practices – or notions about the knowledge-related purposes, methods, and goals of the work in which they are engaged (Sandoval, 2005). We argue that students need to understand and use epistemological considerations for productive and meaningful engagement in scientific practices. For example, developing and revising models that address the mechanism of phenomena lies at the core of the scientific endeavor; thus, considering the degree to which an explanation is mechanistic should guide learners engaged in scientific practice. We termed these epistemological considerations that frame and guide practices "epistemologies in practice" (EIP) (Berland, Schwarz, Kenyon, & Reiser, 2013).

Our prior work has found that students can attend to and productively engage in several epistemological considerations. One such consideration includes students' justification of their knowledge product (such as a model-based explanation) and ranges from students basing their decisions on authoritative claims to basing their decisions on empirical evidence and theory; in short, we label this as the evidence consideration. Another consideration includes students' decisions concerning what kind of answer the knowledge product should provide, and ranges from students providing visible descriptions of phenomena to

providing non-visible causal mechanisms and explanatory processes that explain or predict phenomena; in short, we label this as the mechanism consideration. These, as well as our other considerations, emerged from prior empirical and theoretical work (Schwarz et al., 2012) and share similarities with epistemic criteria in other science education research (Duschl, 2008; Pluta, Chinn, & Duncan, 2011).

Our overall research goal has been to investigate how to support meaningful scientific practices by developing students' EIPs within scientific practices. We aim to determine how and why students' EIP develop or shift over time and to investigate promising pathways for EIP development in scientific practices. Therefore, our research question asks: *How do students' EIPs develop or shift over time and across contexts with respect to the kind of model-based explanations they generate about phenomena?* To address this question, we analyzed data from two high-performing fifth grade students across several science units. We target the analysis around the kind of model-based explanations students generated particularly with respect to whether or how they attended to mechanism, an aspect that played an important role for both students. Mechanism, or causal explanatory processes, are critical in science and for model development and use (Braaten & Windschitl, 2011; Russ, Scherr, Hammer, & Mikeska, 2008). We also analyzed students' reflective talk about their work to examine their notions of what is important in a model-based explanation and how that may have impacted or framed their engagement in the practice.

Method

To determine how students' epistemologies in practice developed or shifted, we interviewed a cohort of ten fifth grade students from four classrooms in 2011-2012 and a second cohort of fifteen fifth grade students from four classrooms in 2012-2013 as they participated in several model-based science units over time. These units engaged students in iteratively developing and revising scientific models that addressed how and why the phenomena of evaporation, condensation, and light occur. All students attended a suburban public elementary school in the Midwest. In this paper, we describe our analysis of data from one student in the 2011-2012 cohort (LS) and one student from the 2012-2013 cohort (JS) who shared the same fifth grade teacher in subsequent years. Both were academically high-performing female students of European-American ethnicity and were chosen for this analysis because they were highly reflective and articulate ten-year-olds. We also chose these students because they made some relatively clear and significant shifts in their EIPs over time, which was often not as visible in other fifth grade students' data. To determine how their EIPs developed over time, we interviewed LS three times (post-evaporation, post-condensation, post-light) in the 2011-2012 school year and three times (pre-chemistry, mid-chemistry, post-chemistry) in the 2012-2013 school year; we interviewed JS four times (pre-evaporation, post-condensation, pre-light, post-light) in the 2012-2013 year.

During the semi-structured interviews, we asked students to describe their (1) models and model-based explanations generated in class, (2) rationales for developing and revising those models, (3) development and application of models and model-based explanations in new contexts, and (4) reflection on this process. For instance, during the condensation interview we asked questions such as: "What did you want your final condensation model to show?" "How does your model answer the question 'How and why do liquids sometimes appear on cold surfaces over time?'" "Looking back at your initial condensation model, what were some important changes that you made and why?" and "Can you use your revised condensation model to explain the phenomenon of how and why rain forms?" As such, the interview elicited information about students' reflections on their model-based explanations and the process of modeling in class, as well as their own practices developing and using model-based explanations.

Table 1 shows the coding rubric for the mechanism EIP consideration we used to code the students' interview responses. The mechanism coding rubric was developed from our prior work (Schwarz et al., 2012) and addresses the kind of answer the knowledge product (i.e., model-based explanation) provides. Similar to Braaten and Windschitl's (2011) framework of ambitious practice for explanation, we coded students' model-based explanations as ranging from students attending to non-mechanistic details (Level 1), to descriptive accounts (Level 2), to mechanistic explanatory processes (Level 3). Descriptive accounts (Level 2) only address *how* a phenomenon happens, which can be a chronological order of events or the condition(s) under which a phenomenon occurs. In contrast, explanatory processes (Level 3) include a causal relationship or mechanism that also addresses *why* a phenomenon happens. These ordered levels do not necessarily imply a sequential or developmental pathway. Instead, they are meant to capture increasing levels of sophistication with respect to the type of explanation students generate.

It is important to note that we did not code the students' responses based on one or two utterances in the interviews. Rather, we analyzed student talk, as well as their diagrammatic model, throughout the interview. For example, if a student provided a partial explanation of the phenomenon in one response and another partial explanation in another response later during the interview, we combined those responses to determine if s/he produced a full mechanism-based explanatory process. By analyzing the students' interview responses and models holistically, we more accurately characterized their model-based explanations about phenomena.

Table 1: Scoring rubric for the type of model-based explanation.

Level	Students' model-based explanation of the phenomena attends to:
3	a partial or full mechanism-based explanatory process that addresses "why it is happening"
2	a partial or full descriptive accounts (including sequences) that addresses "how it is happening" without a mechanism-based explanatory process
1	details that only focus on visible aspects of the phenomenon

Note: If explanations were scored slightly above/below the identified levels, a positive (+) or negative (–) or sign was added (e.g., 2+, 3–).

In addition to using the scoring rubrics to code students' model-based explanations, we also coded the reflective talk that accompanied their decisions and justifications. Students' reflective talk often reveals their views about the purpose of models/modeling or their notions of what is important in a model or explanation, which is critical to better understanding the epistemologies students leverage within their practice. As such, we analyzed LS's and JS's reflective talk to determine why they attended to particular aspects of their model-based explanations. Our analysis of this talk used an inductive approach informed by grounded theory methodology (Charmaz, 2006) to determine reoccurring themes and approaches students brought to justifying their decisions about their model-based explanations. The authors of this manuscript compared and contrasted their individual analysis of students' themes and approaches to decide what themes best fit the patterns in the evidence. While refining our analysis of students' reflective talk, we found that students had particular approaches and ways of framing what was important to include in a model-based explanation. These themes or approaches aligned with several of Braaten and Windschitl's explanation categories (2011), including views about the purpose of an explanation as a simple causal story or as a justification for a claim or argument.

Prior Results and EIP Analysis for Two Cases

Our prior analysis of interview and written data from the larger sample of fifteen focus students indicated that on average, students made modest improvements in their overall EIPs throughout the year and across curricular contexts. In particular, we found productive shifts with respect to all EIP considerations across the evaporation and condensation unit and a smaller shift to the end of the subsequent light unit. Our prior analysis also suggested that the nature of instruction and students' attending to particular EIP considerations impacted how their EIPs developed over time (Schwarz, et al., 2013). Additional analysis of students' embedded assessments from a larger sample of 113 students in a separate study indicated that they significantly improved their attention to mechanism from the evaporation unit to the condensation unit, and modestly improved their attention to mechanism from the condensation unit to the light unit.

In ongoing work to determine how and why such changes might have occurred, we then analyzed the written and interview data from a range of students in greater depth. Our analysis indicated that some students foregrounded heuristics or strategies which seemed to function as lenses, or potential scaffolds, in using their EIP in future contexts. In this study, we report on how two students, LS and JS, did so in particular and non-identical ways. We illustrate their shifts as well as their heuristics or strategies through excerpts of their interview transcripts and diagrammatic models over time. Such data are challenging to present concisely because they span across time and are embodied in particular contexts with specific meanings. Nonetheless, we show how this evidence illustrates visible patterns that contribute to our understanding of how LS and JS developed their EIPs over time with respect to the kind of model-based explanations they generated.

Case 1: LS – Making Models That Explain How and Why

LS was a student from the 2011-2012 fifth grade cohort whom we followed into sixth grade in 2012-2013. In October 2011, we interviewed LS mid-way through the evaporation and condensation unit. We asked LS about the changes she made between her first and second diagrammatic models of evaporation. When asked what she was trying to figure out with her [revised] model (Figure 1), she responded that she was "trying to figure out evaporation and how much time it takes to evaporate." In this instance, LS talked about how long it took for evaporation to occur as a descriptive, rather than mechanistic (i.e., causal), account. When asked if she could talk about the changes she made between the two models, she stated, "I didn't have any temperatures or how and why's [in my first model], so I decided to do how and why [in my revised model]…" While it appears from this reflection that LS may be attending to more causal aspects of the phenomena ('why') in the revised model, she referred to the 'why' in a descriptive manner, indicating that her revisions came from knowing "what actually happens from doing experiments" and ideas from the simulations. When asked how the written descriptions in her revised model improved her model and the reasons she included them, LS responded that "it gives more of an explanation about what happens and more evidence of what I did." At this point, LS's ideas

about her model were primarily focused on capturing what happened from the experiments and the computer simulation in support of her model, rather than capturing how or why they occurred.

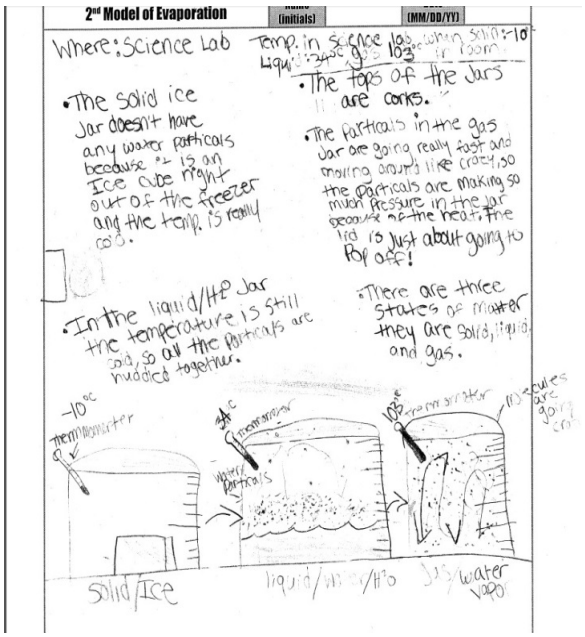


Figure 1. LS's Revised Model of Evaporation

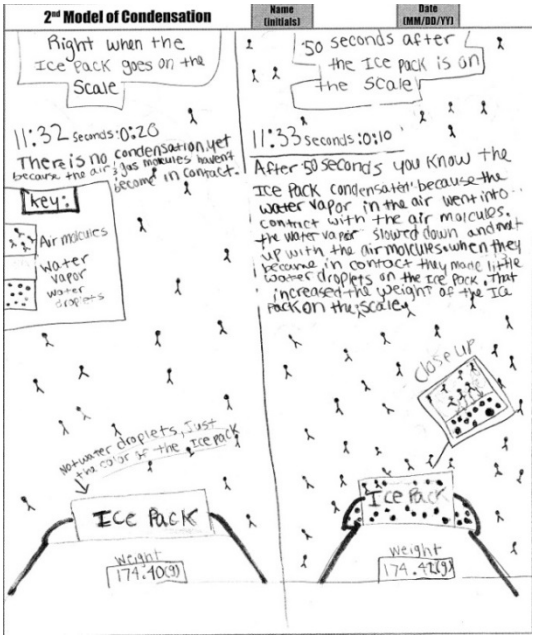


Figure 2. LS's Revised Model of Condensation

Analysis of LS's revised model of evaporation (Figure 1) supports this interpretation and indicates that she focused on the states of matter (something often emphasized in school) rather than the process of evaporation. Although her model was predominantly descriptive (using temperature, etc.), she included details at the molecular level about particles. LS appeared to be documenting information about evaporation without necessarily attending to how or why evaporation was occurring. As such, her model-based explanation and justification was scored at a Level 2 because it provides a descriptive account (sequence of steps) that addresses "how evaporation is happening" without providing a mechanism-based explanatory process (such as water particles spreading out into the air when liquid water comes into contact with air).

When LS was interviewed at the end of the condensation unit in December 2011, the nature of her talk and model showed an increased level of sophistication and orientation towards mechanism, including her notions about model-based explanations (see Figure 2). When asked if she could explain how and why condensation happened in the written question, LS said it was because "the air molecules get cold. … [and the] water molecules [are] coming into contact with the air molecules and making little droplets on the ice pack." This statement, along with her model (in which she wrote "…water vapor in the air went into contact with air molecules. The water vapor slowed down…") and reflective talk, indicated that she provided a partial mechanism for how condensation happens (water vapor comes in contact with cold air, and water vapor slowed down to become water droplets). Her reflective talk showed an increasing emphasis in addressing "how and why" questions to explain her model. For example, when LS was asked what she was trying to figure out with the models she created, her response was that she wanted to know "How and why condensation and evaporation happens." Similarly, when asked if she could talk about the changes she made between the two models, she pointed out that she "…didn't add how and why in this [initial model]…" Her sense of mechanism also became somewhat more sophisticated when she included invisible particles to convince people how the phenomenon happens. When asked which activity addressed this change, LS responded that "…the experiments and probably from the simulations, too, how molecules move and at what temperature and what pace and stuff." Nonetheless, when asked what her group thought was important in a model, she said that they needed "…to have a key and an explanation and evidence and to have people be convinced about what we were doing." At this point in the year, we see that LS is still focused on "what's happening" because she included the actual experiment into her models. At the same time, she also started to recognize the importance of "how and why" when shifting from "what's happening" to "how's it happening" at the molecular level. This resulted in our scoring LS as a low Level 3 for this EIP consideration.

In April 2012, LS finished a short modeling-based unit on the nature of light and was interviewed about her understanding of light and her modeling experiences. Within that post-interview, we again saw that her performance and ideas about mechanism and notions of explanation developed further. For example, when asked, "Could you please describe your final model?" She replied:

> "…So right here, the light source is coming to the person's eyes and also the light source is coming to the cup and the image of the cup is traveling and reflecting to the eye. So I also did reflecting and it says how you see the cup…. It says that the cup is shiny and smooth so [the light] reflects off of it and bounces off. And I said that the light source travels to the cup and an image of the cup travels to the person's eye so that she can see. For scattering, the light bounces off the hard, rough wood table and the light travels and bounces in different directions, making light shine everywhere in the room. … We see things because there's a light pathway …. Scattering happens because the surface is rough, unlike reflection where the surface has to be smooth."

This excerpt illustrates that LS is capable of identifying the reason why people see the cup. For instance, she said that she can see the cup because the light pathway travels from the light source to the cup, which is then reflected to her eyes; this is as an explanatory process. This response is different from a typical fifth grade descriptive account in which a student might say that we could see the cup because of the light, or because the person faces the object. In this excerpt, LS also explained the difference between reflection and scattering by using the same explanatory process, a light pathway, and further discussing how non-visible light moves by bouncing off in different directions. LS describes a fully-mechanistic explanatory process, which is why we scored her response at a high Level 3. Thus, we see LS's model-based explanations and justifications with respect to mechanism become more sophisticated across the year from the evaporation unit to the light unit.

Concurrently, LS's reflective talk indicated that she still continued to think about how and why phenomena happen. When asked what she wanted to figure out with these models, she replied that she "wanted to figure out how and why these things happen…how and why we see things." When asked if she had any goals for herself [in constructing and revising models], she stated, "Well, when we did evaporation and condensation with this, I didn't really know how to make models as well. *So I sort of have a goal set that I would make better explanations for how and why things happen, so I tried to add that into my model as much as I could.*" Soon thereafter, she continued with "If you don't know the how and why, you can't really explain other situations…" which shows her emphasis on the generality of the mechanism.

We interviewed LS during the first unit on chemistry in sixth grade to determine how her EIPs might have changed with a different teacher and new unit. Interestingly, we saw evidence of similar themes in her responses to those from the prior year. For example, we continued to see her emphasis on addressing the "how and why" questions in the pre-chemistry interview. For example, when questioned about why she was asked to draw the model, she recalled the models she had drawn in fifth grade and noted that a model "…helps you know how and why things happen." When asked what she wanted her model to show, LS responded that she "wanted it to show how it happens … because [the teacher] never told us to do a how and why."

In addition to the continued focus by LS from the previous to the present year on addressing the "how and why" questions during the pre-chemistry interview, LS also demonstrated her knowledge of how her model could help her learn. LS said, "…especially when we do more and more because you know more about things and then you can look back at your older models and see if you missed anything. Or when you're revising it, you can look back at your old ones." Similar to the pre-chemistry interview, LS demonstrated her continued focus on explaining "how and why" in the mid-chemistry interview.

In LS's final interview after the end of the chemistry unit, she explained how and why someone could smell an odor by giving a fully-mechanistic (Level 3) "how and why" response. She explained that how the air molecules "are just moving around and once the smell goes, the odor molecules, after they evaporate from the object, they travel around the air together and they will eventually get to your nose. That's why they're moving around all different directions." See Figure 3 for LS's corresponding model. Later in the interview, she added, "If it's more of a hot room they have more energy and they spread out and move faster. If it was a cold room I think they have less energy and they don't move as fast."

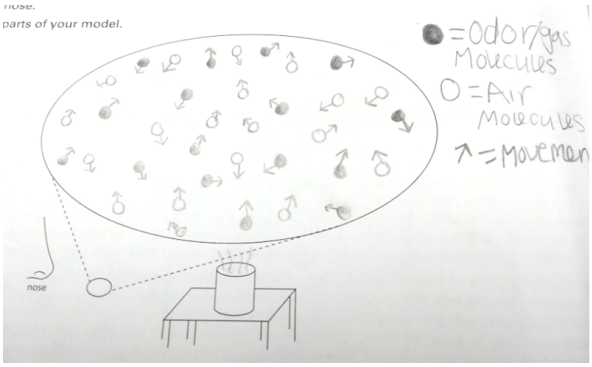


Figure 3. LS's Post Chemistry Model

In summary, we saw LS shift her level of talk and performances, as related to the EIP mechanism consideration, throughout her fifth and into her sixth grade science classes. Her first interview in fifth grade indicated that she foregrounded descriptive and detailed accounts of evaporation (and condensation, to some degree). Over time, we saw LS shift to the idea that models and their explanations should address how and why phenomena occur through her development of models and explanations that provided more mechanistic accounts in the light and chemistry units. In these units, she included non-visible components of the phenomena and described explanatory processes that captured the causal mechanisms involved. As she stated several times in her interview, LS's notions about what counts as an explanation or model (to address "how" and "why") may have guided her in seeking those causal or mechanistic aspects for her models and explanations.

Case 2: JS – Detailed Mechanisms and Evidence to Justify Explanations

JS is a student from the 2012-2013 fifth grade cohort. At the beginning of the year, JS began her pre-evaporation interview with a few ideas about how and why evaporation occurs. When asked what question she could answer with her model, she responded, "I wanted to explain evaporation. I think [the model] could use a little bit more, but I'm not sure exactly what to add yet." When describing her model, she stated, "I drew the lines [in my model] to show like air and it's supposed to show how contact with the air over time makes the water evaporate. I'm not sure exactly how that works." JS stated what she knew about evaporation (by describing that contact with the air makes it happen) but mentioned that she was not sure how it worked. This response was scored a Level 2 for mechanism because she provided a descriptive account with a non-visible component, but did not provide a theory (even if incorrect) to explain why this happened (such as the movement of water into the air).

At the end of the evaporation and condensation unit, we see evidence that JS began to develop a robust sense for mechanism and the importance of empirical evidence. In particular, she seemed to acquire the idea that a convincing causal account for the interactions of components was of paramount importance. JS showed this in a number of ways. First, when asked what question she was trying to answer with her model, she replied, "Well, basically I was trying to answer how and why condensation happens and how it works." Her response illustrates how she was already potentially attuned to mechanism in her language. More detail about what she meant came later in the interview when she asked other questions such as if the various criteria in her model were important: JS responded, "Yeah, they're definitely all important … for mechanism - what's the point of a model if it doesn't explain how it's happening? That's the whole point of it. And evidence [information] sort of shows people that it is possible." She continued on by adding, "For evaporation [the teacher] said that [the water particles] are sort of attracted to the air molecules, but *she didn't say exactly how that worked or anything*." From her response to this question and others, JS seemed to use the term "evidence" to mean "proof" or "information" in the form of a detailed causal mechanism for how the phenomenon occurs. When she was then asked if she though that she had evidence in her model, she replied, "As much as I know. [My model] says that [the water particles] slow down *but it doesn't say why they slow down. [Our teacher] never told us why they slow down, besides that they're getting near something cold*. …it's kind of hard to come by because at this grade they give you some evidence [information] but they don't tell you the whole thing." When asked how she would convince someone that her model was correct, she stated, "Well, I wanted to convince them this is how it happens, but like I said, you need evidence [information]. And I didn't get a lot of evidence [information]. I knew that it happened because I saw it in experiments, but I wasn't sure exactly how it worked…"

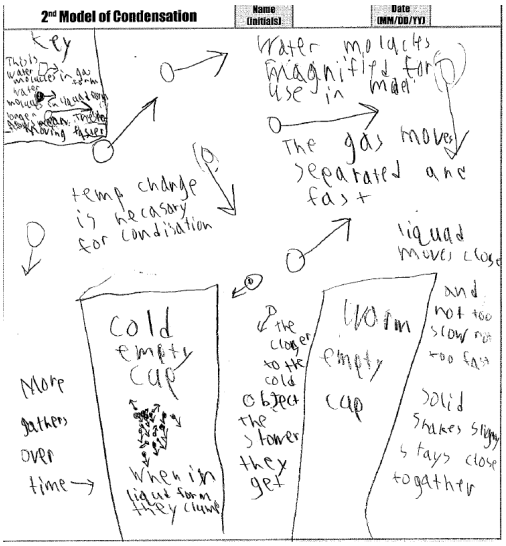


Figure 4. JS's Final Model of Condensation

Overall, her responses indicated that JS prioritized details and accuracy for how condensation occurs. From her model (seen in Figure 4) and in her talk, with respect to condensation, JS described how condensation occurred "as [the particles of water vapor] slow down, they stuck together because they're getting close and closer because they're all slowing down." JS also added that the phenomenon made her think "…[about] the ice pack… I thought that the humidity would go up [around the condensing ice pack] because I thought that the ice from it would end up evaporating. But when [the humidity level] went down, it made me realize that [the ice pack is] pulling the moisture out of the air. Even if there's more water there, if it's cold it's not going to be as much moisture in the air." In summary, JS seemed inclined towards focusing on the detailed causal accounts of condensation, consistent with a Level 3 mechanism score. Her tendency to refine her ideas seemed to focus on comparing alternative ideas to the empirical data collected in class.

When JS started the mini light unit at the end of the year, she began with the same analytical stance on mechanism and evidence that we noted in the evaporation and condensation unit. In other words, JS seemed to use her EIP related to mechanism - her sense for what an explanation involves - as a lens to orient her to the new subject matter. For example, when asked how her model explains how people see, she replied, "The human's eyes can see because they've got some light source. I'm not sure if the light source actually bounces off this or it's just somewhere from some that might have bounced off that and then bounced off a wall or something. But it's a pretty safe guess to say it bounced off this. So once light gets to the eye, the eye – I'm not sure quite how – uses the light to get an image that the brain then processes." While she was not sure exactly how the light entered the eye from the light source, she was willing to think about possible mechanisms and possibilities for how that could have happened (e.g., the light source bounced and eventually gets to an eye to create an image that the brain processes.)

Based on her extensive ideas about how and why light travels, JS seemed to have a fully-developed version of a light model by the end of the unit. Not only did she gain content knowledge about the nature of the mechanism, but she also sought and advanced her ideas about mechanism across contexts. When asked if she could explain what was happening in the model, JS talked about how the light source emits light such that "most of the light gets over here, but some of it hits some bump or groove or something and bounces into a shadow…. And then the light that hit the wall and the little bit of light that hit the shadow bounces back and some of the light that bounces back will hit the eyes to let us see." This excerpt illustrates a causal account for light scattering and reflecting using ideas such as light hitting groves and bouncing back and forth following a pathway. JS seemed to reflect more generally about transferring her analytical stance in explaining phenomena when she said, "If you're actually thinking about the way the world works, you can't look at something and not think of it like you think of everything else." Overall, JS appeared to have developed some ways of thinking about how phenomena occur (using detailed mechanistic explanations of phenomena) that may be carrying forward across contexts.

In summary, we saw JS advance her model-based explanations and reflective talk related to the EIP mechanism consideration throughout the two modeling units in fifth grade. Her first interview highlighted her concern that there was much to understand about the world that she did not know about. In her second interview, JS began to blur the boundaries between mechanism and evidence by her attempts to provide a convincing explanation for how and why something happens. By the end of the year, we saw JS consistently using a highly analytical approach in trying to understand mechanisms of how phenomena occur, even at the beginning of new units, using evidence she thought was convincing. Her notions about having a convincing and detailed model-based explanation may have guided her to seek those causal or mechanistic aspects to make sure they were consistent with empirical evidence in order to address alternative arguments.

Discussion and Implications

While there has been skepticism as to whether students can build their epistemologies in practices across conceptual contexts, our analysis of data from LS and JS indicates that some students can make significant progress constructing more mechanistic explanations of phenomena over time and across subject matter contexts as they considered the kind of model-based explanations they sought to generate. We found that LS focused on creating models and explanations that addressed how and why phenomena occurred. This pathway is consistent with her framing of the "explanation as simple causation" (Braaten & Windschitl, 2011) in which the goal and purpose of the model-based explanation is to provide a causal account of the phenomena. We found that JS focused on creating model-based explanations that were consistent with the phenomena and addressed the "why" question in a highly detailed manner. This pathway is consistent with her framing of "explanation as justification" (Braaten & Windschitl, 2011) in which the goal of the explanation is to justify the causal account with detailed mechanisms and evidence. Both cases illustrate that while the students experienced similar instructional approaches by the same teacher, the students used different pathways to develop their epistemologies in practice by leveraging different resources to frame their endeavors and to advance their work. The students' reflective language indicates that their individual framing about what counts as a model-based explanation may have foregrounded approaches that led them on their unique pathways across subject areas.

There are several important implications of this work. First, this research begins to elaborate possible pathways of students' developing practices by illustrating the importance of framing goals for explanation that can impact how students navigate their practices and develop strategies and approaches. This is important for better supporting students over time through curriculum and instruction. Curriculum can be designed to help teachers recognize and attend to student and scientific goals of the knowledge products as well as how to make those goals explicit for aligning classroom work with those goals. Understanding pathways can also help teachers identify and support EIP development for particular students. Additionally, this research deepens our understanding of how and why scientific practices bridge across contexts in productive ways. Our results show that learners can leverage and carry epistemological considerations of practice across content areas through heuristics and approaches, and potentially use them productively in those new areas. Our case studies indicate that these heuristics and approaches may be due in part to the students' foregrounding of particular epistemologies in practice and by their framing of the endeavor.

We return to our argument that engaging students in scientific practice is not adequate for scientific literacy. Students also need support in attending to epistemological considerations across curricular and instructional contexts in order to make practice meaningful and productive. As such, we advocate that EIPs should be considered through the thoughtful design of curriculum materials and instruction that align with discussions in which classrooms address questions such as, "Why are we doing what we're doing? What's our goal? How will we know whether we've gotten there?"

References

Berland, L., Schwarz, C., Kenyon, L., & Reiser, B. (2013, April). *Epistemologies in practice: Making scientific practices epistemologically-rich for students.* Paper presented at the annual meeting for the American Educational Research Association: San Francisco, CA.

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, *95*(4), 639–669. doi:10.1002/sce.20449

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks, CA: Pine Forge Press.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*(1), 268–291. doi:10.3102/0091732X07309371

Ford, M. J., & Forman, E. A. (2006). Redefining disciplinary learning in classroom contexts. *Review of Research in Education*, *30*, 1–32. doi:10.2307/4129768

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. (H. R. Quinn, Ed.). Washington, DC: The National Academies Press.

National Research Council. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*(5), 486–511.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525. doi:10.1002/sce.20264

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634–656.

Schwarz, C., Reiser, B., Acher, A., Kenyon, L., & Fortus, D. (2012). *MoDeLS: Challenges in defining a learning progression for scientific modeling.* In A. Alonzo & A. Gotwals (Eds.) Learning Progressions in Science (LeaPS) (pp. 101-137). Boston, MA: Sense Publishers.

Schwarz, C., Reiser, B., Davis, B., Kenyon, L, Acher, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Designing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal for Research in Science Teaching, 46*(6), 632-654.

Schwarz, C. V., Akcaoglu, M., Ke, L., Zhan, L. (2013, April). *Fifth grade students' engagement in modeling practice across content areas: What epistemologies in practice change over time and how?* Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, UK: Cambridge University Press.

Acknowledgments

We thank the Practices research group, the participating teachers and students, and the anonymous reviewers for their feedback. This research was funded by the National Science Foundation grants DRL1020316 and DGE 0707432. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors.

MOOCs: A Perspective from the Learning Sciences

Michael Eisenberg and Gerhard Fischer
Department of Computer Science and Institute of Cognitive Science
University of Colorado, Boulder USA
duck@cs.colorado.edu, gerhard@cs.colorado.edu

Abstract: Massive Open Online Courses (MOOCs) have-with astonishing rapidity–become a formidable presence in global education. Such courses have obvious strengths in their ability to convey (usually lecture-based) content at extremely low cost to widespread, and often underserved, student populations. At the same time, MOOCs, for the time being at least, reflect traditional (and often contested) values in education: the default assumption for a MOOC is that a teacher or professor will present facts or interpretations, record these as video or slide presentations, and convey them to an extensive audience of (usually individual) students. This paper addresses the default assumptions of MOOCs in the light of two decades of research in the learning sciences, focusing on themes of collaborative work, embodied cognition, and both the limitations and opportunities of learning analytics. With these themes in mind, we suggest paths for research and exploration into the next generation of MOOC design.

Keywords: MOOCs, collaborative learning, embodied and extended cognition, isolation problem, shop-oriented MOOCs, learning analytics

Introduction

To anyone who has been attentive to global education over the past several years, the importance and explosive growth of *Massive Open Online Courses* (MOOCs) is both striking and undeniable [Scientific-American, 2013]. Arguably, MOOCs–or, more generally, the advent of online education–can be viewed as the most urgent subject for discussion in the learning sciences for the coming decade. There are, unsurprisingly, dozens of controversies and unresolved issues [Daniel, 2012] regarding such courses: (1) whether they are destined to grow in impact or (alternatively) whether their importance is overrated, and a symptom of media "hype"; (2) whether they represent an improved or cheapened educational experience; (3) what effects they may have on the structure and future of residential universities; (4) how their impact might vary across dimensions of geography, of demographics (are MOOCs as important for elementary education as for undergraduate education?), and of discipline (are MOOCs more easily designed for technical material?).

The state of the art in online education is both relatively short-lived, and fluid; generalizations made about a "typical MOOC" today are liable to be out of date in a matter of weeks or months. Recent technological history should be a chastening experience in this respect: it is not hard to find statements made about the nature of personal computing in (say) 1977, or about the importance of the internet in (say) 1995, that upon reflection embarrass their authors. Even with this caveat in mind, however, it is important for those in the learning sciences both to reflect upon and to attempt to shape the development of large-scale online education; and we have no choice but to make the attempt in real time, as events unfold. For the purposes of this paper, we begin by noting that despite the novelty of the transmission channel, in many respects MOOCs tend to represent a highly traditional portrait of education, in which a teacher or professor delivers lecture-based material to a (relatively passive) audience of students. This portrait of a "typical course" is hardly unusual in most universities, as can be seen by attending a standard introductory-level course in calculus, or chemistry, or Western history. At the same time, this traditional portrait has been challenged in numerous respects by various communities in the learning sciences over the past two decades. Such notions as collaborative learning [Koschmann, 1996], embodied and extended cognition [Dourish, 2001], and constructivist learning [Johnson, 2010] are staples of learning sciences research, but in many ways they run against the grain of what, for now, could be called a traditional lecture-course structure for course design. In residential universities, the limitations of lecture courses are often addressed through creative strategies such as "flipped classrooms", intensive recitations or tutorials, and study groups; these strategies not infrequently employ MOOCs or online lectures as components. Still, in contrast, a stand-alone MOOC outside a residential setting (based as it is on "pure" traditional lectures) might find these strategies tricky to implement. In other words, while both MOOCs and residential universities might employ the traditional lecture-to-passive-students format for education, the residential settings are able to experiment with alternatives or additions that are harder to implement in a purely online format.

 © ISLS

Background

This paper is intended to "take a step back" and discuss the current limitations and potential progress of MOOC design from the perspective of the learning sciences. It is thus intended as a complement, or counterweight, to many other accounts that discuss the subject from an *economic* perspective (scalability, productivity, low cost) [Bowen, 2013] or the *technological* perspective (platforms supporting large numbers of students in online environments, enrichment components such as forums, peer-to-peer support, automatic grading) [Kolowich, 2013]. Our perspective is derived from our own research over the past two decades in creating computational environments for learning, incorporating notions such as collaborative, embodied, and extended cognition [Arias *et al*., 2001; dePaula *et al*., 2001; Eisenberg & Eisenberg, 1999]. Our interest is also intensified by a newly-begun (arguably quixotic) project in which one of the authors is attempting to create an unorthodox MOOC based as much on the format of a reading seminar as a lecture course. With these experiences as both foundation and motivation, we discuss *the challenges to incorporating learning sciences research within MOOC design*, and we outline several *promising directions* for integrating the work of the learning sciences with the near-term future of large-scale online education.

Collaborative Learning in the Context of MOOCs

One of the notable features of MOOCs is their flexibility from the standpoint of the student: in general, an online course can be taken from any Web-enabled site around the globe, at any time of day, in any setting. This advantage of flexibility is accompanied by a concomitant problem of isolation: many, perhaps most, students of MOOCs are placed in a situation in which they are taking a course without the immediate presence of a roomful of classmates. As a result, it is tricky or effortful for MOOC designers to incorporate elements of collaboration or group work in a MOOC: one can use online tools such as student blogs, chat sites, shared documents, and so forth for these purposes, but they require additional effort on the part of both the professor and students beyond the natural affordances of the physical classroom.

A significant thread in learning sciences research over the years has in fact focused on the characteristics of successful (or sometimes, unsuccessful) collaborative learning. Without attempting a summary of the extensive literature, suffice it to say that there are situations in which problem solving can be facilitated by conversation [Roschelle 1992]; in the same vein, the literature of self-efficacy theory in education takes note of the potentially positive role of "vicarious experiences" in which a student benefits from watching the techniques of a successful colleague [Bandura 1997, p. 86 ff.]. Beyond these considerations, group work allows for creativity from the professor's standpoint: in many disciplines there are challenging projects or assignments that are most naturally undertaken not by a single student, but rather by a small group or team [Fischer, 2011]. Collectively, these considerations suggest that there is an essential tension between the default structure of MOOCs and the potential advantages of student collaboration. (For fascinating reflections along these general lines from a "MOOC practitioner", see also [Downes 2012, p. 503 ff.])

All of this suggests that there is a strong incentive for applying the lessons of learning sciences research to the "isolation problem" of MOOCs. Some of the natural avenues for work are incremental: finding modest, realistic means of enhancing Web-based tools for collaborative student work and real-time conversations over a distance. Just to take a representative example: it is well-known that the simple act of pointing to an object as a focus of conversation plays an important role in human learning. (Indeed, Tomasello [1999] among others argues that pointing is a fundamental act differentiating human culture from that of other primates. For a recent study of the role of pointing drawn from mathematics education, see [Alibali & Nathan, 2012].) Pointing is, of course, a natural act for two humans in conversation working at a lab desk, but it can be an unwieldy operation for students conversing at a distance, particularly when the object being referenced is a physical entity, not on a computer screen. Conceivably, then, a learning-sciences-motivated research effort could be undertaken in which a user (call her "Alice") can point to a physical object present in the space of a second user (call him "Bob") through the use of a remotely-operated robotic hand. Roughly, the image here is that Alice can use some type of glove-like input device at her site to manipulate a robotic pointing output device at Bob's desk, enabling her to point directly to (say) a misplaced wire on a breadboard or a troublesome spot in a mechanical construction.

Other possible research projects might be a bit more futuristic or far-fetched, but nonetheless plausible. Conceivably, for example, individuals could advertise their services not as local *tutors* for course material, but as local *colleagues*: people whose job it is to learn alongside a student and act as a "sounding board" for conversations and lab partner for projects. Such people could, in effect, specialize in joining others to take courses [Fischer, 2013]; the skills required of a successful "co-learner" might, over time, come to be seen as comparable to (but distinct from) those of a successful teacher or student. Still another possibility, along these same lines, would be to design artificial, computer-based colleagues (again, in contrast to the tradition of work in computer-based tutors [Anderson *et al.*, 1995]); that is, the intent would be not to create computer-based presentations of existing material, but rather to create (where achievable) computer-based "learning partners" [Choua *et al*., 2003] whose role would be similar to those of the hypothetical human specialists just mentioned. The purpose of an "artificial colleague", then, would be to provide at least some of the benefits of real-time

collaboration (e.g., by assisting in problem-solving, note-taking, finding resources related to lecture material, and so forth). This, too, would represent a potentially interesting novel direction for the learning sciences and artificial intelligence: indeed, in some ways, it may in fact be more achievable (and more useful) to create a working artificial *student* to accompany a human learner, rather than an artificial *teacher*.

Weaving Embodied and Extended Cognition into Online Education

Perhaps the most vexing and fundamental problem associated with MOOCs (and arguably with online education in general) is the overwhelmingly "virtual", screen-based nature of the presentation channel. Screen-based education seems to lend itself fairly naturally to the recording of lectures, along with slides, text, audio, and video; it can also, to a fair degree, accommodate text-based exchange of information (as in student blogs or email), and audio or video conferencing. All of this, in combination, makes for a powerful medium that can be tailored to the presentation of a wide variety of course material: one could imagine well-done introductory courses in history, philosophy, computer science, and English literature developed along these lines. In general, the types of courses that one associates with large lecture rooms on college campuses ("Calculus 101") seem to lend themselves reasonably smoothly to translation into MOOC form.

A recurring theme heard from critics of MOOCs in higher education, however, is that there are many courses that *do not* fit especially well into this mold. What about, for instance, lab courses in physics, or organic chemistry, or cell biology? All the equipment, staffing, materials, safety precautions, and so forth are parts of *physical settings* for learning experimental science; what happens to those elements in the MOOC format? Can one possibly learn experimental natural science exclusively through an online medium? For that matter, what about field trips in domains such as geography, or paleontology, or archaeology? The very essence of MOOC education is precisely that it is *indifferent* to settings; one can take a MOOC anywhere, at any time, all over the globe. But a physical geography course takes on a different character if it is held in (say) Buenos Aires, or Minneapolis, or Hong Kong. To raise yet another issue, what about the example (so often mentioned by college faculty) of seminars or small-scale conversation? Is it possible to convey, through online form, anything like the value of a group of committed students sitting around a table [Arias *et al*., 2001] discussing a work of literature, or a philosophical argument?

The difficulty of bridging the gap between MOOCs and the wide range of models (beyond the lecture hall) of undergraduate education connects with several running themes and traditions in learning sciences research. One natural place to start here is with the twinned themes of *embodied and extended cognition* [Dourish, 2001; Menary, 2010; Clark 1997] In the context of this discussion the two themes, though distinct, have shared pedagogical implications. Extended cognition characterizes cognition and learning in the presence of external artifacts–for our purposes, such artifacts might include lab equipment, shop tools, sports equipment, and so forth. To the extent that learning chemistry takes place through the use of pipettes, burettes, and separation funnels, it would be difficult to teach "true" chemistry exclusively via video, in the absence of such equipment. Embodied cognition characterizes cognition and learning as an activity of the entire body (often with a disproportionate focus on the hand), which again implies the importance of pedagogical tools and settings. To the extent that learning how to observe an infant in the laboratory is a matter of posture, tone, and "body knowledge", or to the extent that learning to shoot a basketball is a matter of information beyond the purely verbal, it seems unlikely that video will be a sufficiently expressive medium; this is why we have lab apprenticeships and sports coaches.

If the very real, but very specific, advantages of MOOCs are to be enriched beyond the context of lecture hall material, it seems imperative that learning scientists work with designers to incorporate the findings of extended and embodied cognition research into novel artifacts and settings. We might begin this process by reflecting on some exemplary disciplines. Is it possible to teach a student to conduct a safe laboratory experiment exclusively via online communication? Or to operate a drill press? Or to shoot a basketball? Or to give a haircut? If not, why not?

On the *"extended cognition"* side of these questions, we might ask (among other questions) just how much it is possible to replicate or simulate the presence of physical tools in MOOC-accessible settings. Perhaps it would be feasible to augment online communication with desktop 3D printing so that at least some implements could be output in physical form at the student's site. Admittedly, this would be only a very partial, or stopgap, means of addressing the absence of laboratory equipment: desktop 3D printers at present output objects in limited materials (typically plastic) and of small size (no more than about a cubic foot). Still, the ability of students and staff to communicate via transmission of physical objects would be an important extension to a strictly screen-based medium. The role of learning sciences research in such an effort would be to understand just *how* important, or meaningful, such an extension might be, and to direct design research whose goal is to move beyond the current limitations of educational desktop 3D printing.

There are other ways of addressing the "laboratory gap" in MOOC education beyond the purely technological. The growing "maker movement" [Anderson, 2012] of amateur, hobbyist, and student construction is increasingly associated with a burgeoning network of physical sites–"makerspaces", or "hackerspaces"–in

which participants (usually for a modest fee) can make use of physical materials and equipment for their own projects. Combining the presence and growth of settings such as these might allow for extended versions of *"shop-oriented MOOCs"*: for instance, one might create a course that includes "shop videos" whose intent is to be shown and followed in a local makerspace setting. This would represent a strategy of *infrastructural* rather than purely technological design–creating accessible educational settings that are distinct from schools or universities (or even school laboratories) [National-Research-Council, 2009]. The role of the learning sciences here would be to understand and enhance the learning process in such settings. Much as learning scientists have studied processes of social learning, language-making, and domain-oriented dialogue in schools (for particularly provocative foundational discussions in a vast literature, see for instance [Brown, 1992] and [Schoenfeld, 1985]), there would be a necessity to understand the adaptation of these ideas to makerspaces, with their greater emphasis on physical construction.

More important, at least in the near term, it would be useful for the learning sciences to pay increased attention to the educational dimension of the maker movement itself, as it currently exists. For those children and undergraduate-age students who participate in "Maker Faires" and similar events, what sorts of projects are they doing? What content are they learning (and how does that content contrast with that of traditional schools or universities)? What are the means by which learning takes place, and how can (or can't) these be integrated with MOOC design? By understanding the educational dimensions of the burgeoning "DIY" culture, the learning sciences can direct attention to precisely those aspects of "hands-on" learning that for the present seem to elude MOOC-based education.

As to the *"embodied cognition"* side of this discussion, there are again a variety of approaches-technological and infrastructural-that could enhance MOOC design. On the one hand, we could see a vast expansion of innovative wearable or body-based devices for communicating with computational systems. A MOOC student who can make use of a data glove, or a device like the soon-to-be-marketed MYO (with input derived from large-scale arm movements; https://www.thalmic.com/en/myo/), or sensors placed on arms and legs, might be able to take effective online courses in violin, painting, or how to throw a curveball. There are vast unexplored regions for the learning sciences in integrating an understanding of such devices into cognitive models and educational design; indeed, the increasing popularity of MOOCs might well spur a new round of input techniques and body sensing that could only be accomplished with the assistance of a rich theory of embodied cognition. We could imagine the use of these new-generation devices being employed to create MOOCs for (currently) unexpected subject matter: an online system, equipped with the right sort of input device, might be able to help a student wield a paintbrush, or conduct an orchestra, or titrate a solution, or grip a tennis racket.

There is an infrastructural side to the "embodied cognition" dimension of this discussion as well, focusing on the integration, within MOOC education, of settings for full-body interaction. Conceivably, MOOCs might be tailored to work with local science museums, with participating playgrounds, or with national parks. For example, a MOOC on the subject of geology might include a "general component", not specific to any location, and consisting of standard online material; and it might be augmented with a localized component tailored to the proximity of specific public parks or spaces. Again, this represents new questions for the learning sciences to explore in order to make courses of this type meaningful for the students: what makes for an educationally useful (or memorable) field trip? How should one structure these trips? (Are individual visits to, e.g., public national parks worthwhile, or should these visits be done in groups, and, if so, of what size?) How can one tailor the structure of these courses to international settings [Lewin, 2013]?

The essential point in this discussion is that, without the participation of researchers who understand the role of gesture, physical surroundings, and full-body movement in learning, MOOCs are destined to be highly cramped, constrained media of education. Designers of online courses need the learning sciences to create worthwhile experiences; and learning scientists can use the MOOC phenomenon as a source of new research and design for exploring embodied and extended cognition.

Learning Analytics and the Challenges to a Truly Worthwhile Online Education

One of the most interesting aspects of MOOC education is that it affords researchers access to large data sets of student choices, behavior, and responses. *Learning Analytics* (with its own society (http://www.solaresearch.org/) and its own conference series (http://www.solaresearch.org/events/lak/)) focuses on the way in which researchers can study the order in which students understand particular concepts or material, the patterns of errors that they make in answering questions, the portions of lectures that they watch (or re-watch) in their studies, and so forth. These questions have antecedents in the literature of computer education (e.g.: in intelligent tutoring systems), but what is novel about MOOCs is the sheer volume and variety of the data sets: it is now possible to study the learning behavior of thousands of students on specific questions, where before it would have been effortful to derive similar data sets for dozens of subjects. The large numbers of students also allow statistical analyses of students' behavior "broken out" by factors such as age, geographic region, socioeconomic status, prior educational experience, and so forth.

The growth of MOOCs, and the associated growth of techniques for learning analytics, poses interesting philosophical challenges for the learning sciences. We can begin by asking whether this data is in fact what people interested in learning really *need*? To what purpose will the data be put? For instance, if the goal of education is to ensure that students make fewer errors in answering questions, then learning analytics can clearly be put to that purpose. Certainly "making fewer errors" is *one* goal of education–everyone would prefer to answer questions correctly–but is it an important or primary goal? Likewise, we might use the patterns of a student's answers to speed up or slow down the pace of a course. No one would object to this practice, but is it the foundational issue facing the learning sciences?

To put the matter another way: even assuming that all the currently envisioned benefits of learning analytics are eventually realized in full, it is not at all clear how much students' educational experiences will be improved. This is an empirical question, but it is at least arguable that a better-paced course with fewer errors made in answering questions will make only a modest impact on the educational issues that matter most to human beings.

There is actually a *cautionary* note here for the learning sciences. It is conceivable that the growth of MOOCs will result in a narrowing of the portrait of the learner necessary for educational improvement. That is, the "learning sciences" might become increasingly the study of very large patterns of data in answering questions or viewing video segments; such study would increase insight in some dimensions while suppressing it in others. We might find that the learning sciences pay less attention to the idiosyncratic, necessarily individual patterns of interest growth [Csikszentmihalyi, 1990]; or that studying anthropological factors in education (family dynamics, peer pressure among teenagers, how close friendships influence the development of academic interests, and so on) is increasingly "under the radar" of the sorts of large-scale data collected by MOOC enterprises. If we want to answer more biographical questions about education-such as: What is it about certain settings that makes them inspirational for children? What kinds of experiences do children have with handheld mathematical puzzles? Why do adolescents sometimes admire peers who actively spurn intellectual activity? Why don't some children ever read for pleasure?-then we have to energetically pursue a type of learning science that goes well beyond the content of learning analytics, at least as that field is currently envisioned.

There might be still other ways in which this tension between the "big data" style of MOOC research and the "biographical", or "narrative" type of questions described here could be alleviated. One might, for example, design novel types of online education created expressly to be addressed to "best-friend pairs"; that is, the "student" unit in this scenario would be a pair of close friends working together. Or MOOCs could be designed to be used by a family, or group of buddies at a lunch table, or the population of a neighborhood. In scenarios of this sort, we could imagine the powerful techniques of learning analytics being turned to investigate the behaviors of friends or families or close-knit groups. Large quantities of data could be gathered on how families interact with course material, or how peer groups support (or suppress) learning, or how friendships evolve in the presence of intellectual challenge; the data gathered in this way might be unavoidably limited, but nonetheless point toward new questions for ethnographic observation. MOOCs could thus be tools for (at least some measure of) anthropological inquiry, and not merely a tool for massive-scale abstract cognitive modeling. This discussion returns us to some of the ideas of the earlier discussion on collaboration, and "going beyond the individual student"; the techniques described above for encouraging collaboration could also be augmented by large-scale means of data collection and analysis.

Some Closing Reflections: the Learning Sciences in Changing Times

Learning science, as an academic enterprise, is approximately 20-25 years old; though young, as scientific disciplines go, it already has an identifiable history. The origins of the learning sciences, as we understand them today, are traceable in cognitive science, developmental psychology, and computer science in the second half of the twentieth century. It is fair to say that the learning sciences have, throughout this period, been closely interwoven with an understanding of the technology of the moment. The very ideas of cognitive modeling, intelligent tutoring, identifying problem-solving strategies, and the like emerge from metaphors closely associated with computation and information processing; the tools and innovations most closely associated with learning sciences research tend likewise to be computational systems and languages.

The technology of the moment, in education, seems increasingly to lean away from classroom instruction and toward online instruction, particularly in higher education. In some ways this is a promising shift–it might herald democratized access to high-quality education, and a vast increase in numbers and diversity of committed students. At the same time, there are presently clear limitations of MOOCs as means of human communication. One could of course point out (and many do, persuasively) that there are substantial limitations to conventional classrooms or lecture halls; but the rapid advent of MOOCs compels us to reconsider all these limitations with fresh eyes. The learning sciences, by focusing on several of these limitations–of isolated individuals as students, of screens as media of embodied learning, of learning analytics as portrait of data collection–can help us re-imagine a technological landscape for education that moves beyond the emerging

portrait of higher education: (1) Techniques that foster *in-person conversation and collaboration*; (2) artifacts that increase (e.g., through gesture, or direct physical fabrication) the *expressiveness of online communication*; (3) infrastructure that respects the processes of *learning "through the hands" or "through the body*"; and (4) an attention to *anthropological and individual biographical data* in addition to the massive population-level data collection associated with MOOCs-all these are suddenly urgent tasks for the learning sciences, where they may have seemed relatively peripheral before.

A profitable way to approach these questions might be to step back and re-examine fundamental questions associated with our field in the light of new technology. When we speak of "the learning sciences", what portrait of learning do we have? What comes to mind? One interesting division is between "learning *about*" and "learning *to be*" [Brown, 2005]. To learn *about* is to focus on the accumulation of intellectual capital, organized into a curriculum that stresses the communication of culturally central theories, facts, and skills [Hirsch, 1988]. A curriculum of this sort is most naturally structured as a sequence (ideally, a fine-grained sequence) of educational objectives; and the methodology for "learning about" is frequently linked to the introductory lecture-hall style of delivery, augmented with readings, problems, and tests. In short, when we think about "Calculus 101" or "American History 101" we are drawn into a portrait of "learning about" as an objective for education.

In contrast, learning *to be* is to focus less on teaching mathematics, physics, or history, and more on what it means and takes to be a mathematician, a physicist, or a historian (or, for that matter, a "Wikipedian", skier, or surfer). Education in the "learning-to-be" mold centers on engaging students in personally meaningful problems, encouraging teachers to model problem-solving activities in front of or alongside their students rather than lecturing (cf. [Schoenfeld, 1985]), and enculturating students into communities of practice or professional behavior (cf. [Nash & Shaffer 2011]) via techniques such as "legitimate peripheral participation" [Lave & Wenger, 1991]. In the context of this discussion, the current limitations of MOOCs (and the need for learning science research) could be characterized as enhancing the ability of new modes of education to foster "learning to be" through conversation, varied communities, physical activity, participatory settings, and anthropological methods.

The objectives of "learning about" and "learning to be"–like other twinned objectives in educational discussion (e.g., between "informal" and "formal" learning, or between "symbolic" and "kinesthetic" understanding)- represent *antinomies* [Bruner, 1996]: pairs of complementary truths, each worthwhile to be pursued in different contexts, but also presenting learners and educators with tensions and contradictions. Antinomies of this kind have woven their way, historically, into enduring questions in the learning sciences: for instance, whether computers should be designed as independent stand-alone "intelligent tutors" or (by contrast) as "intelligent assistants", or "expressive tools of communication", or as some mixture of these and other artifacts (for an early example of these debates, see [Taylor 1980]). Over time, we are likely to come to view MOOCs and their offspring through the lenses of similar antinomies. Are we after *uniform, standard, globalized* high-quality education (a chance to take the very best "Calculus 101" course for every student in the world)? Or are we after *highly personalized, idiosyncratic* pathways of learning [Collins *et al*., 2009] (a chance for a student with any particular interest–say, in the study of European ferns, or the geography of the planet Mercury, or Viking ship design–to fashion a unique educational experience geared toward their own preferences)? To take another example: are we after an *inexpensive* educational infrastructure (in which students can easily afford at least a minimal education, and in which the resources associated with residential universities are scaled back) or an *expanded* infrastructure (in which online education is complemented not only by residential universities, but by more widespread resources for construction, laboratory work, or field trips)?

It is in the nature of such educational antinomies that they elicit multiple responses, depending on the material to be learned, the students, the setting, and many other factors. The essential goal of the learning sciences in the face of new technology is to identify the various sides of the antinomies latent in the technology; once identified, we can use the technology in an informed way, research its role in learning, and design alternative or complementary technologies that mitigate the problems of one-sidedness. Where MOOCs (and online education) are concerned, we are faced with a potentially powerful technology–especially powerful in economic and political terms. It is plausible that not only higher education, but also K-12 education, professional training, vocational education, and graduate work may be increasingly defined by the characteristics that we see, in embryonic form, in MOOCs today. If the science of learning is to serve our ends as human beings, we must pitch our theories to the types of learning that really matter to us, and address our designs to the creation of artifacts that enable us to fulfill our aims.

References

Alibali, M. and Nathan, M. (2012) Embodiment in Mathematics Teaching and Learning: Evidence from Learners' and Teachers' Gestures. *Journal of the Learning Sciences* **21**(2), pp. 247-286.

Anderson, C. (2012) *Makers - the New Industrial Revolution,* Crown Business, New York.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995) "Cognitive Tutors: Lessons Learned," *The Journal of the Learning Sciences,* 4(2), pp. 167-207.

Arias, E. G., Eden, H., Fischer, G., Gorman, A., & Scharff, E. (2001) "Transcending the Individual Human Mind—Creating Shared Understanding through Collaborative Design." In J. M. Carroll (Ed.), *Human-Computer Interaction in the New Millennium,* ACM Press, New York, pp. 347-372.

Bandura, A. (1997) *Self-Efficacy.* New York: W. H. Freeman.

Bowen, W. G. (2013) *Higher Education in the Digital Age,* Princeton University Press, Princeton.

Brown, A. (1992) Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, **2**(2), pp. 141-178.

Brown, J. S. (2005) *New Learning Environments for the 21st Century,* available at http://www.johnseelybrown.com/newlearning.pdf.

Bruner, J. (1996) *The Culture of Education,* Harvard University Press, Cambridge, MA.

Choua, C.-Y., Chanb, T.-W., & Lin, C.-J. (2003) "Redefining the Learning Companion: The Past, Present, and Future of Educational Agents," *Computers & Education,* 40, pp. 255-269.

Clark, A. (1997) *Being There.* Cambridge, MA: MIT Press.

Collins, A., Fischer, G., Barron, B., Liu, C., & Spada, H. (2009) "Long-Tail Learning: A Unique Opportunity for CSCL?" In *Proceedings (Vol 2) of CSCL 2009: 8th International Conference on Computer Supported Collaborative Learning, University of the Aegean, Rhodes, Greece*, pp. 22-24. Available at: http://l3d.cs.colorado.edu/~gerhard/papers/2009-CSCL-symposium.pdf.

Csikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience,* HarperCollins Publishers, New York.

Daniel, J. (2012) *Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility,* available at http://www-jime.open.ac.uk/jime/article/viewArticle/2012-18/html.

dePaula, R., Fischer, G., & Ostwald, J. (2001) "Courses as Seeds: Expectations and Realities." In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *Proceedings of the European Conference on Computer-Supported Collaborative Learning*, Maastricht, Netherlands, pp. 494-501.

Dourish, P. (2001) *Where the Action Is — the Foundations of Embodied Interaction,* The MIT Press, Cambridge, MA.

Downes, S. (2012) *Connectivism and Connective Knowledge.* Available at: http://www.downes.ca/files/Connective_Knowledge-19May2012.pdf

Eisenberg, M. and Eisenberg, A. (1999) Middle tech: Blurring the division between high and low tech in education. In Druin, A. (ed.), *The Design of Children's Technology.* San Francisco: Morgan Kaufmann, pp. 244-273.

Fischer, G. (2011) "Understanding, Fostering, and Supporting Cultures of Participation," *ACM Interactions* XVIII.3 (May + June 2011), pp. 42-53.

Fischer, G. (2013) "From Renaissance Scholars to Renaissance Communities: Learning and Education in the 21st Century." In W. Smari, & G. Fox (Eds.), *International Conference on Collaboration Technologies and Systems,* IEEE, San Diego, pp. 13-21. Available at: http://l3d.cs.colorado.edu/~gerhard/papers/2013/CTS.pdf.

Hirsch, E. D. (1988) *Cultural Literacy: What Every American Needs to Know,* Vintage, New York, NY.

Johnson, S., ed. (2010) *Neoconstructivism: the New Science of Cognitive Development*. New York: Oxford University Press.

Kolowich, S. (2013) *With Open Platform, Stanford Seeks to Reclaim MOOC Brand,* available at http://chronicle.com/article/With-Open-Platform-Stanford/142783/.

Koschmann, T. D. (Ed.) (1996) *CSCL: Theory and Practice of an Emerging Paradigm,* Lawrence Erlbaum Associates, Mahwah, NJ.

Lave, J., & Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation,* Cambridge University Press, New York.

Lewin, T. (2013) *U.S. Teams up with Operator of Online Courses to Plan a Global Network,* available at http://www.nytimes.com/2013/11/01/education/us-plans-global-network-of-free-online-courses.html.

Menary, R., ed. (2010) *The Extended Mind.* Cambridge, MA: MIT Press.

Nash, P. and Shaffer, D. (2011) Mentor modeling: the internalization of modeled professional thinking in an epistemic game. *Journal of Computer-Assisted Learning* **27**, pp. 173-189.

National-Research-Council (2009) *Learning Science in Informal Environments — People, Places, and Pursuits,* National Academy Press, Washington, DC.

Roschelle, J. (1992) Learning by collaborating: convergent conceptual change. *Journal of the Learning Sciences,* **2**(3): 235-276.

Schoenfeld, A. (1985) *Mathematical Problem Solving.* San Diego, CA: Academic Press.

Scientific-American (2013) *Learning in the Digital Age (Special Report),* available at http://www.scientificamerican.com/editorial/digital-education/.

Taylor, R., ed. (1980) *The Computer in the School: Tutor, Tool, Tutee*. New York: Teacher's College Press.
Tomasello, M. (1999) *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.

Acknowledgments

The challenges, reflections, ideas, and envisioned developments articulated in this paper are based on (1) the collaborative research of the two authors in technology enhanced learning over the last two decades in the Center for Lifelong Learning & Design (L3D) and its support by the National Science Foundation; (2) the numerous contributions of the researchers and PhD students in L3D; (3) the discussions with colleagues in the learning science community (specifically: Allan Collins, Mark Gross, Sharon Derry, Frank Fischer); and (4) the interactions with researchers who have developed or are developing MOOCs (specifically: Mitchel Resnick, Don Norman, Pierre Dillenbourg, Sriram Sankaranarayanan). This work was also supported in part by the National Science Foundation under grant IIS0856003.

How Interpreters Make Use of Technological Supports in an Interactive Zoo Exhibit

Brian Slattery, Leilah Lyons, Priscilla Jimenez Pazmino, Brenda Lopez Silva, Tom Moher
University of Illinois at Chicago, 1240 W. Harrison St., Chicago, IL 60607
Email: bslatt2@uic.edu, llyons@uic.edu, pjimen5@uic.edu, brendita@uic.edu, moher@uic.edu

Abstract: Informal science institutions are depending more and more on dynamic, technology-enabled exhibits. These engaging firsthand experiences often need to be contextualized so that peripheral audiences are able to learn from them. We investigate how interpreters (a.k.a. docents, explainers) equipped with a tablet support tool (TST) are able to facilitate learning in a dynamic exhibit with high visitor traffic. This TST focused on re-representing user performance in the exhibit, and was designed to help interpreters shape visitor discourse and inquiry. Four cases are presented to illustrate how interpreters orchestrate their collaborative facilitation when a meditational tool is present, and how the epistemological role of TST shifts. We document interpreters successfully using TSTs to coordinate shifts in facilitation based on the state of the interactive exhibit, and using live data representations to connect visitors' personal experiences with exhibit content, demonstrating several potential uses of TSTs as mediational tools for dynamic exhibits.

Background and Introduction

Informal science learning in institutions such as museums, zoos, and science centers is depending more and more on interactive technology and dynamic exhibit content, which is a trend that has been noted by the National Research Council (Bell et al., 2009) as well as the Learning Sciences community (Yoon et al., 2013). But an increased reliance on dynamic, technology-dependent exhibits increases the degree of human-technology interaction while often decreasing the degree of human-human interaction, raising the challenge of how to support groups of learners at these exhibits without disrupting the beneficial social learning that takes place (Heath, vom Lehn, & Osborne, 2005; Hall & Bannon, 2006; Hornecker & Stifter, 2006).

Support for human-human interactions within technology-based exhibits could come from a resource already present in informal learning institutions: interpreters (also known as docents, explainers, or facilitators), who are trained to facilitate learning for large, diverse groups of museum visitors through shared meaning-making around exhibit resources (Tilden & Craig, 1977; Falk & Dierking, 2008; Beck & Cable, 2012). Successful interpreters are expert improvisers: they assess existing visitor interests and knowledge levels, and attempt to build bridges between visitors and the exhibit's content. This becomes more challenging for dynamic digital exhibits, where the state of the exhibit can change moment-to-moment. This suggests that there may be value in a digital tool that can help interpreters remain apprised of these changes to the exhibit's state. There is some research on the value of interpreters appropriating technology to facilitate dynamic exhibits (Hsi, 2008), indicating that there is a potential role that interpreters can play in these new kinds of exhibits. Our question is how to intentionally design technological support tools that can be used by interpreters in their facilitation of dynamic exhibits, especially for the goal of supporting the learning of a large, diverse visitor base that is often peripheral to the central exhibit interactives.

The possible utility of designing technology to support facilitators of collaborative learning has been acknowledged for formal learning environments, especially through the use of mobile tools, which have been designed for both teachers and students (Roschelle & Pea, 2002). Interpreters could be considered the "teachers" of informal learning environments, but they face challenges that are not shared by their formal learning counterparts. Learners in informal spaces vary widely in age and background, so there is no guarantee (or even clear indication) that they have been exposed to particular representations, or ways of thinking. Also, informal institutions are "free-choice" spaces where learners are able to decide what they want to do from moment to moment, so there is variation in how long visitors will spend in any given interaction with an exhibit or interpreter. Thus, it is unclear what lessons can be applied from the design of technological supports for formal learning instruction, as these rely on an expectation of instructor autonomy and control that is not possible for interpreter-facilitated learning. Instead, the design of support technology for interpreters must come from a detailed understanding of its situated use in an informal learning context, as this highlights the consistencies and inconsistencies of practice and enactment that drives the iterative improvement of a design (Bodker, 2009).

Designing for Dynamic Exhibit Interpretation

A successful tool for interpreters is one that would serve as a mediational means for interpreters' interaction with visitors, allowing them to engage in facilitative practices that they would not have normally been able to perform without the tool (Wertsch & Rupert, 1993). In the context of interactive, dynamic exhibits, where

visitors are generating exhibit content through their own actions, support technology could potentially help interpreters to shape the discourse around visitors' contributions, influencing visitors' expectations for learning and interaction (Gutierrez, 1994). It would be helpful to provide interpreters with real-time guidance on how to orchestrate learning for groups of peripheral visitors who are not interacting directly with an exhibit, since the content they can engage with changes based on the actions of visitors who are participating.

Interpreters could also benefit from a support tool that improves their ability to recontextualize exhibit resources for the purpose of improving visitors' meaning-making (Falk & Dierking, 2008). Interpreters are trained to draw analogies and make personal connections between themselves, visitors, and the exhibit, which in dynamic exhibits involves connecting different modalities of exhibit content (e.g. visitor movements translated into real-time graphs of performance data) as well as grappling with variation in the visitor-driven representations that learners are able to make sense of and reason around. In this difficult meaning-making context, a support tool would be beneficial if it allowed interpreters to position exhibit resources epistemologically as accessible sources of inquiry, facilitating visitors' own sense-making of the exhibit (Bezemer & Kress, 2008; Jaipal, 2010).

This work has taken a combined action research and user-centered design approach to exploring how to design meditational tools for interpretation. From the practice-oriented literature on interpretation we can glean some lessons for what makes for successful interpretation experiences, but we must acknowledge that as with any meditational tool, the creation of the tool simultaneously disrupts existing practice while making new forms of practice possible. We investigated interpreter practice by working closely with the interpretation staff at Brookfield Zoo in Chicago, IL. We then tested the prototype *in situ* with actual interpreters, so that we could observe how the tool would be appropriated in an actual use context. Observations made during this *in situ* study provide the cases that this paper presents.

Design of the Exhibit and Support Tool

With this study, we introduced a dynamic, technology-driven exhibit into a zoo environment. This exhibit, *A Mile in My Paws* (Jimenez Pazmino et al., 2013), presented new interpretive challenges to the zoo staff, as it relied on player-generated content, presented a semi-controversial topic (climate change), and used real-time representations (e.g. player-generated line graphs) that are not commonly used in zoos. In the exhibit, only one visitor at a time controls a virtual polar bear with "swimming" and "walking" motion controls to traverse an arctic environment in search of food. *Paws* presents computer generated arctic environments in the past, present, and future, highlighting the reduction in sea ice extent and subsequent increase in polar bears' caloric expenditure (it requires more energy to swim rather than walking across ice).

Our goal was to gather information about how interpretation looks in this context, for the purpose of improving interpreters' practice through iterative design revisions of *Paws* and the TST. However, since we are not interpreters, our understanding of the practice of interpretation needed to come from the community of interpreters that we were working with. One way we approached this understanding was through conversations with Brookfield Zoo's interactive programs manager, who outlined the interpretive training that takes place at the zoo, which ultimately derives from the principles proposed by Tilden & Craig (1977). This approach stresses the importance of conversational dialog between interpreters and visitors, with the interpreter focusing on the visitors' unique interests and background knowledge, the exhibit's learning goals and expected takeaways, and any possible analogies or references that can clarify exhibit content. Additionally, the primary designer of the TST took an action research approach (Noffke, 1997) and spent thirty hours "embedded" as an interpreter, participating in the standard interpretation training given to new hires as well as interpreting at several traditional exhibits with visitor inquiry activities. This helped ground the more abstract theoretical knowledge of interpretive practice with experience enacting these practices, which gave the design team a more situated understanding of interpreters' core concerns and goals. This also informed our later analysis as it allowed us to contextualize differences in TST usage in terms of interpreter practices, and balance our research goals against the real-world demands of interpretation when designing the initial TST prototype.

The prototype TST was made to provide interpreters with resources for contextualizing the *Paws* player's actions in terms of data representations relevant to climate change literacy. These include maps of sea ice extent over time, and graphs illustrating changing trends (in this case, the virtual polar bear's caloric expenditure over time). These would allow interpreters to make connections between exhibit content (both the immediate events going on in *Paws* and the foundational climate change content) and visitors' individual experiences and understanding, which is a facilitative skill that interpreters normally exercise at static exhibits and animal viewing areas. The performance graphs were also intended to provide additional information to peripheral visitors about the level of exertion of the *Paws* player, as relative levels of exertion are typically difficult to judge by observation alone (Rejeski, 1981; Jimenez Pazmino et al., 2013).

We also included "just-in-time" discussion prompts that would appear on the TST based on the state of the player's progress. The prompts presented questions about polar bears, climate change, and human activity that had been gleaned from observations in previous *Paws* pilots and discussions with interpretive staff. These

were designed to aid interpreters in identifying upcoming "teachable moments" (Jimenez Pazmino et al., 2013) that might resonate with events occurring onscreen or drive ongoing discussions with visitors. The TST interface was designed in Unity and implemented on an iPad for the pilot studies at the zoo (see Figure 1).

Study Setting and Procedures

After two controlled pilot studies (Jimenez Pazmino et al., 2013), our team organized a two-day installation of the *Paws* exhibit in Brookfield Zoo's underwater polar bear viewing area. Each day, the exhibit was available for roughly one and a half hours. Although exhibits at Brookfield Zoo are typically staffed by two interpreters working independently, at any given time *Paws* was being facilitated by a slightly larger team of two to five interpreters due to its novelty. These interpreters were mainly part of the "Roving Naturalist" program at Brookfield Zoo, which is a paid seasonal interpretation program made up of college- and middle-aged interpreters, and which focuses on both scheduled and spontaneous discussions with visitors. Each day, one of the interpreters was also drawn from the "Youth Volunteer Corps," an unpaid interpretive program for high-school-aged teenagers. All interpreters involved with *Paws* were already trained in facilitation using artifacts, specimen carts, and other specialty installations, and had been introduced to the idea behind *Paws* and its main learning goals. Only one of the interpreters ("Lorraine") had prior experience facilitating *Paws,* as part of a preplanned talk with a youth summer camp group.

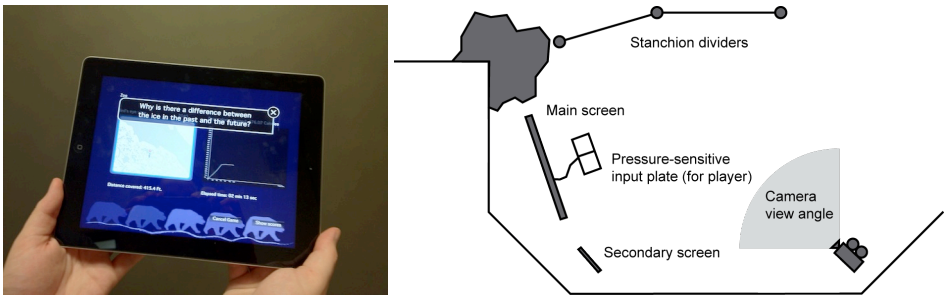


Figure 1: On the left, the main view of the prototype TST, showing progress map, calorie graph, and just-in-time question prompt. On the right, an overview of the elements that make up the *Paws* exhibit.

Interactions between visitors and interpreters were captured through a combination of video and audio recordings. Interpreters were equipped with lavalier microphones that recorded both their speech as well as the speech of visitors in their immediate vicinity. A fixed camera at the back of the exhibit (see Figure 1) provided an overview of *Paws* and captured visitor and interpreter gesture, gaze, stance, and relative positioning. After this implementation, we held two post-hoc design revision meetings that included Lorraine as well as Brookfield Zoo's interpretive programs manager and another interpreter who had received additional training in facilitation for climate change and arctic issues. These meetings were recorded and gave additional insight into the reasoning behind the style of facilitation Brookfield Zoo's interpreters are trained in.

Using Quicktime Pro and ChronoViz, a qualitative data analysis application, the audio and video recordings of the two-day implementation were combined using multiple audio tracks (from the different lavalier microphones). For each day, the researchers identified and marked the spans of time where an interpreter was holding the TST. These timespans were transcribed, with particular attention paid to interactions where the TST was referenced or used by either the interpreter holding it, the visitor, or another interpreter.

The cases below were selected to show instances where interpreters were using the TST as part of their facilitation. When a *Paws* session was in progress (i.e. for almost all of the time recorded), interpreters were continuously conversing with visitors, but not all of these conversations referred to TST resources. This is not surprising, as the TST was still relatively new to this group of interpreters. These cases are therefore meant to illustrate some potential ways that the TST can be incorporated into and mediate interpreters' facilitation.

Case Analyses: Orchestrating Facilitation with the Support Tool

When facilitating *Paws*, the interpreters faced a challenge that was novel in some ways compared to their typical interpretive tasks, which generally are comprised of individual interpreters having extended discussions with families and small groups, or presenting scheduled, lecture-style talks with preplanned content. *Paws*, in comparison, incorporated multiple learning goals related to climate change (e.g. understanding regional change over time, representational literacy, "experiencing" polar bears' struggles) as well as a dynamic simulation that required extended visitor participation. The TST was designed to help interpreters orchestrate *Paws* by presenting multiple representations of the player's progress (location on the map, distance and time meters, and a caloric expenditure graph), but the introduction of the tool also impacted the relationship between the interpreters in the exhibit. In one case on the first day of the implementation, Lorraine accidentally interrupts

Claire's discussion with a group of visitors around the differences in sea ice between the different simulations of the arctic environment:

> <*Claire walks forward towards mixed-age group of visitors around stanchions at the periphery of the exhibit. She is holding the TST in front of her like a screen*>
> **Claire**: So now in 2045 there's gonna be a *lot* more *water* than there will be ice, so she's gonna be swimmin'-- swimming a lot more [*points back to current Paws player*] than our last guest, <previous player> did, in 2010.
> <*Short pause as Claire looks back at current player and visitors look at TST and the current player, while Lorraine approaches the same group of visitors*>
> **Claire**: <*simultaneously*> And that's--
> **Lorraine**: <*simultaneously*> Now have you guys-- sorry--
> **Claire**: Go ahead.
> **Lorraine**: Sorry, have you guys gotten to see the polar bears here? We have two polar bears, do we know-- do you guys know who our polar bears are here?
> <*Lorraine has discussion with visitors about the polar bears at the zoo, Claire turns back to the* Paws *player*>
> **Claire**: <*cheering on the player*> Good job, keep it goin', you're doin' a good job!

Claire begins a discussion about changes in sea ice extent due to climate change, showing the player progress map on the TST to a group of visitors, and connecting this to the differences in effort between two *Paws* players who experienced different years. However, Lorraine accidentally interrupts Claire's discussion during a short lull, and introduces a different conversation about the identities of the resident polar bears at the zoo. Claire decides to turn backwards from the group and join another interpreter in cheering on the player's efforts (which was a common action for many interpreters to engage in when there weren't visitors available to speak to).

This case illustrates the additional attentional demands that the TST creates for interpreters. Typically, joint facilitation of an exhibit (which is common at Brookfield Zoo) involves a pair of interpreters discussing exhibit content with visitors. Interpreters largely use social norms to negotiate conversational turn-taking, but the introduction of TST complicates this arrangement. It creates a situation where visitors' silent observation (or manipulation) of a screen is an additional "relationship" present in the social space. This can make it difficult for interpreters to judge when they should change the conversational focus (since the TST creates another source of stimuli that interpreters must monitor), as well as complicating visitors' attention (as the TST can play the role of another "participant" in the conversation). It is especially critical that interpreters are able to negotiate these new attentional challenges when using the TST.

In the above case, the visitors were mainly at the periphery of the exhibit, which may have made it additionally difficult for Claire and Lorraine to accurately judge the visitors' engagement with the exhibit. In the following case from day two, Lorraine and Audrey—facilitating for a group of visitors who, in this case, were gathered at the center of the exhibit—used the TST to support their joint facilitation based on conceptual and temporal divisions between different *Paws* topics, at the point where one run of the simulation was ending:

> <*Lorraine and Audrey are standing on either side of the main screen. One child is controlling the bear ("Player") while the other three are pantomiming along by mimicking the player's swimming and walking motions ("Pantomimers 1-3"). This behavior is also being modeled by Lorraine who is in the middle of an extended interaction with the visitors. Audrey is holding the TST and watching their progress*>
> **Lorraine**: […] So we gotta get to that seal. <*Speaking to the pantomimers*> Are you helpin' her swim? Are you helpin' her swim?
> **Pantomimer 1**: Why-- why--
> **Pantomimer 2**: Yeah I am.
> **Pantomimer 1**: Why does she-- need to eat the seal?
> **Player**: Cuz I'm hungryyy!
> **Lorraine**: It's because, is-- is she a person? No, she's a what?
> **Children**: <*in unison*> Polar bear.
> **Lorraine**: And polar bear-- and what--
> **Audrey**: <*interrupting*> Alright you're almost close!
> **Lorraine**: Oh she's almost there!
> **Pantomimer 1**: Yeah!
> <*The player reaches the seal*, *which is the goal of the game*>
> **Lorraine**: She got it, yay, she eats her seal, it's delicious!
> **Audrey**: Alright so I'm--

Lorraine: Good job, good job.
Audrey: [*looking down at TST*] Alright, so in two minutes and eighteen seconds, you burned four hundred and eighty nine—[*turns to group and points to screen*] fifty two calories. So if you were a polar bear [*points to screen, turning TST away from group*] and you were tryin' to get that food, that's how much calories you would've burned in that little time.
Player: Is that good?
Audrey: That's a lot, that's a lot of calories. And if they burn a lot of calories, that means that their body's gonna go lower and lower and they're gonna have to find more food.
<*One of the pantomimers says something inaudible*>
Audrey: So do you guys think that it's possible to help out polar bears that live over there, do you think it's possible that we can do something here to help them out? Like what? What do you guys think we can do?

As the player is nearing the end of their run, Lorraine guides the discussion around a few topics: the identity of the player as a virtual polar bear, the differences between running and swimming, and that seals are eaten by bears for energy. Audrey interrupts this conversation when she sees the player is approaching the seal on the TST map. Upon reaching the seal, Audrey reads out the amount of calories that the virtual bear spent, and has a back-and-forth with the player about the meaning of bears spending energy to find food, and that bears might not be able to recoup their caloric losses if they have to travel far. She then shifts the discussion towards climate change mitigation actions that visitors can engage in, asking the children for ideas on how their actions can affect the arctic environment.

In this case, the TST mediates Audrey's interaction with the visitors, as she plays the role of shifting the visitors' focus towards complex topics of caloric expenditure and climate change mitigation. Here, the TST is used to expand on ideas introduced in an abstract manner during gameplay (e.g. energy usage), which these visitors were already attending to as a group. Audrey builds on the ideas that Lorraine introduced, situating them in the immediate experience that the visitors had just shared (e.g. What does the player's performance mean relative to polar bears' need for energy?). Rather than using the TST to continue the discussion led by Lorraine, Audrey uses the TST to move from playing *Paws*, to reflecting on the larger topic of climate change mitigation that forms the context for the entire exhibit. Audrey's use of the TST was also skillful because she recognized the correct time to interrupt Lorraine's interaction. If Audrey had tried to shift the visitors' focus to climate change mitigation before the game had ended, she would not have been able to make use of the final "calories expended" value displayed on the TST to motivate the interest of the visitors, and instead would have had to rely on a less immediately relevant argument. When facilitating dynamic player-driven exhibits, it is critical that interpreters have a keen sense of the appropriate time to introduce new ideas into a discussion. Covering instructional content "just-in-time" allows for interpreters to make use of coherences between the player experience and the behavior of the simulation (Crowley & Galco, 2001; Jimenez Pazmino et al., 2013). Although the first case shows that the TST can introduce new challenges for how interpreters coordinate their joint facilitation, it is also able to provide representations—in the second case, the map of player progress—that aid interpreters in judging when is the optimal time to step in and engage in facilitation around a new topic.

Case Analyses: Shifting the Support Tool's Role in Inquiry

One of the central practices emphasized in Brookfield Zoo's interpretive training is connecting the personal experiences of visitors with the instructional content at exhibits, which is often removed from visitors' everyday understanding of biology or ecology. While making content relevant is difficult and can be a challenge in formal learning environments, it is additionally so in informal learning environments where visitors—especially zoo visitors—are very diverse in age, interest, background knowledge, and group composition (e.g. parents with teenage children, elderly couples, elementary school groups). This means that interpreters have to be adept at having back-and-forth conversations with visitors. Lecturing didactically or engaging in rigidly structured interactions such as the common Initiate-Response-Evaluate classroom pattern (Mehan, 1979) would not offer serendipitous information about a learner's interests or knowledge. This information is critical to interpreters as they are trained to seek out unexpected ways of making exhibit content relevant to visitors.

Often, interpreters will rely on personal anecdotes and comparisons between their own and visitors' lives as a way to initiate conversations, build rapport, and establish analogies that can be used to explain unfamiliar topics. They are also trained in using props, especially animal remains (nicknamed "skulls and skins"), to serve the aforementioned conversational functions. In a sense, the TST is one such prop, as it provides at-hand visuals that can be used to engage visitors. However, the TST differs in that it mainly displays re-representations of the digital exhibit (the progress map and calorie graph). This allows TSTs to play an analogical role affording a different point of access to exhibit content, similar to interpreters' personal anecdotes. But since the TST is built around data representations, the interpreters perceived it as less accessible or attention-grabbing for visitors compared to a personal story or connection. Thus there is a risk that during an

interaction with visitors, the interpreter might have difficulty establishing the epistemological role of the TST (what function it is supposed to serve as a source of knowledge for supporting collaborative learning).

During the first day of the implementation, Kristina (who was the one teenaged YVC interpreter at the exhibit) used the TST to shift an "off-topic" conversation towards the *Paws* exhibit's core learning goal:

> *<An adult visitor ("Visitor") who had been cheering on the player initiates a conversation with Kristina, who is holding the TST, about how to make the game even more "embodied">*
> **Visitor**: What would be funny is if you had a squirt bottle and a sprayer.
> **Kristina**: It would-- that would be-- that's actually kind of a good idea.
> **Visitor**: [*laughs*]
> **Kristina**: I don't know if we're allowed to do that, we might have to look in-- that would be fun, just when you hit the water.
> **Visitor**: Just when she hits the water.
> **Kristina**: [*laughs*] Now you swim, yeah.
> **Visitor**: Yeah, that would be fun. [*laughs*]
> *<Kristina angles the TST towards the visitor>*
> **Kristina**: And you see if you pick the year like-- if you pick the projected year of 2045 *<Visitor looks down at TST>* you'd get squirted a lot more often.
> **Visitor**: Yeah.
> **Kristina**: More water in the area. *<Visitor looks up and nods>*

Although the conversation begins with a seemingly minor remark about an addition to the exhibit facilitation, Kristina subtly turns the conversation towards the central exhibit topic of change in sea ice and polar bear energy usage over time. The visitor may not have even noticed the TST when she initiated the conversation, but Kristina was able to include it in their conversation fluidly because it was already in her hands. Without the TST, Kristina would not have had a means of connecting future sea ice extent, the player's experience, and the squirt bottle idea directly, assuming she would have even chosen to make that conversational move at all. In this case, the TST is being positioned as a source of information about a potential experience ("if you pick…2045 you'd get squirted a lot more often") but also as a bridge from visitors' personal contributions to exhibit content. Having a ready means of drawing connections between the unpredictable statements of visitors and the topic of a nearby exhibit is something that interpreters value. The TST gives them a resource that they can rely on to turn a conversation towards the topics they are trained to discuss, and that they judge as valuable for learners.

Later in the first day, another interpreter, James, has an extended interaction with a family group as the three young siblings ("Eli", "Aaron", "Adam") take turns playing *Paws* one after another (Figure 2 below). Most families with multiple siblings would only allow one child to play, since each run took about five minutes, which is a lengthy stay in the context of typical zoo exhibits. However, James interacted with the group for fifteen minutes while the other interpreters at the exhibit worked with the player and other visitor groups. This gave James multiple opportunities to position the TST as a tool for inquiry.

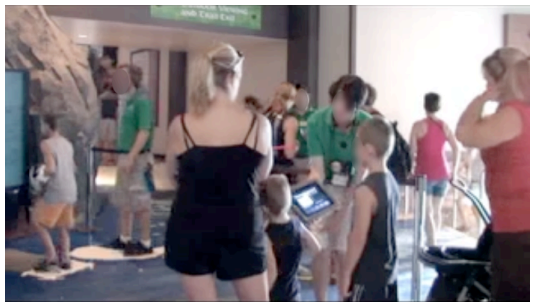


Figure 2. James (green shirt, middle) shows the TST to Eli and Aaron while their sibling plays *Paws*.

> *<James standing with Eli, Aaron, and their mother. Aaron has just finished playing* Paws *and is James' focus>*
> **James**: [*leans down to show Aaron the TST*] Alrighty, so now you can kinda see what you had to travel, you see that distance from the blue dot to the red star? That's what you traveled. But you had, did you think, did you think there was more ice or more water?
> **Parent**: < ? > faster too on ice.
> **Aaron**: Wa- ice.
> **James**: Right, that's right. So, now did you notice, was it harder to swim or was it ease- harder to walk?
> **Aaron**: Harder to swim.

James: Exactly, it was harder to swim. So you can kinda see how that's gonna happen on this little chart. 'Cause graphs are awesome, they tell you a little bit about everything.
<Later in the conversation>
James: [*kneels down and shows TST to Eli and Aaron, who both look at it*] Now! On this graph! Where do you think he was swimming? When less energy was used, or more energy? [*pointing at different segments of the line graph*]
Eli: More.
James: More? Yeah. So you can kinda take a look at this graph, and it's gonna kinda show you how over time, more energy is used when you're swimming. So that's proof right there, that you're not crazy, you do get more hungry after you swim.

Here, James is positioning the TST—specifically the caloric expenditure graph and map—as a source of information, in particular as a way to find out more about what the player experienced and how it changed over time. This differs from the more didactic stance that interpreters take when presenting exhibit content in scripted talks. Even though the conversation is relatively one-sided (part of this likely being due to the age of the children), James repeatedly positions the TST as the source of his observations about the game state, by making reference to the graph as proof that "you're not crazy" and that more energy is used while swimming. As Brookfield Zoo has been increasingly incorporating inquiry activities into facilitation, this interpretive practice aligns well with their institutional goals. By using the TST and its data representations as an accessible means of providing evidence for questions, rather than just being proof of an interpreter's assertions, it becomes available for inquiry driven by both visitors and interpreters. Despite the fact that these questions ("Was it harder to swim or was it harder to walk") were stated by an interpreter, and were asked because they lead to a particular desired conclusion (i.e. the visitor realizing that swimming is harder), the TST was being positioned as providing access to the answer, which is a critical epistemological move for instructors (Jaipal, 2010).

Interpreters using the TST had a range of resources that they could draw on as they interacted with visitors, but they did not use all of these resources in the same way, or in the ways that we designed. For instance, the "just-in-time" discussion prompts were designed to aid interpreters in connecting the ongoing activity at the exhibit with *Paws* learning goals, but they were not used to drive visitor inquiry. Rather, the map and graph were more frequently used for this purpose, even though the interpreters had to specifically clarify the connection of those resources to the activity at hand during discussions with visitors. Interpreters exercised their own *ad hoc* judgments of how best to facilitate visitor inquiry with TST and *Paws* resources (e.g. James' question about Aaron's experiences as a player), rather than relying on the pre-designed contextual prompts. Both Kristina and James were able to connect the *Paws* players' experiences with the re-representation of those experiences on the TST, indicating their ability to position the TST as a means for visitors to gather evidence about questions the interpreters had posed. We still believe "just-in-time" notifications can be helpful, but would be better used for non-inquiry practices, such as supporting interpreters' situational awareness and their ability to collaboratively shift visitors' attention and discourse, rather than initiating and supporting visitor inquiry.

Discussion and Conclusions

Even with little training in how to facilitate exhibits with a handheld tool, the TST was able to support interpreters by allowing for facilitative practices that they did not previously have access to. Interpreters were able to use the resources and awareness provided by the TST to coordinate collaborative facilitation of the exhibit. The interpreters transitioned between topics based on the state of the interactive exhibit, requiring interpreters to make use of their own training as well as situational awareness provided by the TST to make these shifts at appropriate times. Interpreters were also able to use live data representations as a way to establish connections between visitors' personal comments and exhibit content, as well as for positioning exhibit resources as available for collaborative meaning-making. Interpreters facilitated inquiry in this way by generating questions around TST resources that allowed those resources to be used by visitors as evidence. The fact that lightly-trained interpreters were able to enact these practices suggests that TSTs hold much promise, certainly as a mediational tool in interactive exhibits like *Paws,* but possibly in more traditional exhibits as well.

This work raises a number of questions about interpreter practice, especially about the decision-making process that interpreters use to guide their interaction with other interpreters and visitors. Why did interpreters choose to discuss particular topics at the times they did? How and when should the TST be strategically employed to guide visitor inquiry? Investigating these questions would also contribute to literature and training around interpreter-supported visitor inquiry (Garibay et al., 2010), by introducing the potential for technological tools to support and facilitate inquiry in informal learning settings. These tools provide interpreters with ways to facilitate STEM skills (e.g. situating abstract data representations) with visitors, including with younger learners that have limited representational fluency.

Our next revision of the TST is centered on improving interpreters' situational awareness and providing them with a wider range of resources. This involves keeping easy access to key representations such

as the map and graph that interpreters used consistently to organize their facilitation, while also presenting and sorting additional resources (e.g. sea ice seasonal variation, polar bear habitat shifts, caloric expenditure of analogous human activities) based on their relevance to the current state of the simulation. This is designed to allow interpreters to keep track of the player's progress, identify potential connections they can illustrate for the peripheral visitor audience, and have greater flexibility of resources that can be incorporated into discussions with visitors. The introduction of a TST for interpreters doesn't just improve on existing facilitative practices, but has the potential to support qualitatively different forms of facilitation. This is possible when interpreters recognize the ways in which the TST can change the contextual landscape of an exhibit, and can change the kinds of conversations that interpreters have with visitors.

References

Beck, L., & Cable, T. (2012). *The gifts of interpretation: Fifteen guiding principles for interpreting nature and culture*: Sagamore.

Bell, P., Lewenstein, B., Shouse, A., & Feder, M. (2009). *Learning science in informal environments: People, places, and pursuits*. National Research Council.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication, 25*, 166-195.

Bodker, S. (2009). A human activity approach to user interfaces. *Human-Computer Interaction, 4*(3), 171-195.

Crowley, K., & Galco, J. (2001). Everyday activity and the development of scientific thinking. In K. Crowley, C. Shchunn & T. Okada (Eds.), *Designing for science: Implications from everyday, classroom, and professional settings*. Mahwah, NJ: Erlbaum.

Falk, J., & Dierking, L. (2008). Enhancing visitor interaction and learning with mobile technologies. In L. Tallon & K. Walker (Eds.), *Digital technologies and the museum experience: Handheld guides and other media*: AltaMira Press.

Garibay, C., Martin, L., Rubin, A., & Wright, T. (2010). Math in Zoos and Aquariums: The Evolution of a Professional Development Workshop.

Gutierrez, K. (1994). How talk, context, and script shape contexts for learning: A cross-case comparison of journal sharing. *Linguistics and Education, 5*, 335-365.

Hall, T., & Bannon, L. (2006). Designing ubiquitous computing to enhance children's learning in museums. *Journal of Computer Assisted Learning, 22*(4), 231-243.

Heath, C., vom Lehn, D., & Osborne, J. (2005). Interaction and interactives: collaboration and participation with computer-based exhibits. *Public Understanding of Science, 14*(1), 91-101.

Hornecker, E. & Stifter, M. (2006). *Learning from interactive museum installations about interaction design for public settings*. Paper presented at the 2006 Computer-Human Interaction SIG of the Human Factors and Ergonomics Society of Australia (OzCHI).

Hsi, S. (2008). Designing for mobile visitor engagement. In L. Tallon & K. Walker (Eds.), *Digital technologies and the museum experience: Handheld guides and other media* (pp. 125-146): AltaMira Press.

Jimenez Pazmino, P., Lopez Silva, B., Slattery, B., & Lyons, L. (2013). Teachable mo[bil]ment: Capitalizing on teachable moments with mobile technology in zoos. In *Extended abstracts of the 2013 Conference on Human Factors in Computing Systems (CHI EA 2013)*. Paris, France.

Jaipal, K. (2010). Meaning making through multiple modalities in a biology classroom: A multimodal semiotics discourse analysis. *Science Education, 94*, 48-72.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.

Noffke, S. (1997). Professional, personal, and political dimensions of action research. *Review of Research in Education, 22*, 305-343.

Rejeski, W. (1981). The Perception of Exertion: A Psychosociological Integration. *Journal of Sport Psychology, 4*, 305-320.

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1*(1), 145-168.

Tilden, F., & Craig, R. (1977). *Interpreting our heritage*: University of North Carolina Press Chapel Hill.

Wertsch, J., & Rupert, L. (1993). The authority of cultural tools in a sociocultural approach to mediated agency. *Cognition and Instruction, 11*(3/4), 227-239.

Yoon, S., Quintana, C., Lyons, L., Perry, J., Osterweil, S., & Lindgren, R. (2013). *Promises and perils of using digital tools in informal science learning environments: Design considerations for learning*. Paper presented at the 10th Int'l Conference on Computer Supported Collaborative Learning (CSCL).

Acknowledgments

This work was supported by NSF CCEP-I Grant 1043284.

Becoming an Activist: Learning the Politics and Performances of Youth Activism Through Legitimate Peripheral Participation

Joe Curnow, Ontario Institute for Studies in Education, University of Toronto
252 Bloor St West, 7th Floor, Toronto, ON M5S 1V6 joe.curnow@mail.utoronto.ca

Abstract: Bringing situated learning theory to social movement theory, this paper examines the ways young adults engaged in social justice activities become activists. As novices within United Students for Fair Trade, students reported not identifying with activism, yet through their immersion in USFT's community of practice and their increasing participation in the dominant practices of the community, particularly the facilitation techniques, norms, and rituals, they came to identify and be identified as activists. This study highlights the value of situated learning and community of practice theory for social movements while demonstrating legitimate peripheral participation.

Learning and Becoming in Social Movements

How do students become activists? This study analyzes the engagement of university students involved in social justice activities as they adopt the identity of activists through their work in social movements. Using United Students for Fair Trade (USFT) as a case study, I trace the identity development of youth activists and analyze how they learn through legitimate peripheral participation in an activist community of practice. Examining learning and identity in social movements enables us to see how people become involved, gain experience, and become active change-makers in their communities and around the world.

Learning Sciences approaches have brought much value to concerted studies of cognition in context, particularly within schooling contexts. They allow researchers to see and examine in detail the learning processes that are enacted and maintained within communities. This type of analysis is valuable and should be extended to other contexts. My interest is in bringing the same intensity of analysis and attention to context to social movements so that we can understand how learning enables people to participate in social change work and develop activist identities. Studies of learning in social movements have increased within the past decade (Hall and Turay, 2006), and yet few rely on sociocultural theories of learning. These conceptual frameworks would bring value to adult education and social movement theory, as they are better able to describe and theorize the ways that individuals and collectives co-produce meaning, generate strategies for action, and mobilize their learning into social movements.

My research questions ask how USFT students became activists and how their identities and learning opportunities were produced through involvement in an activist community of practice. Given the theme of this year's ICLS, "Learning and Becoming in Practice," this analysis is particularly apt, as it centres on the ways that young adults engaged in social justice work learned, and how that learning shifted their identities through their activism. Participants' social movement practices actively constructed the community and enabled them to move from dis-identifying with the label of activist to actively embracing and enacting the dominant practices of activism.

Theoretical Approach: Situated Learning and Legitimate Peripheral Participation in an Activist Community of Practice

As a highly influential theory of learning, situated learning, with key concepts legitimate peripheral participation and communities of practice, (Lave and Wenger, 1991; Wenger, 1998) represents an important stronghold of theorization that offers an alternative to acquisition models of learning. Lave and Wenger argue that learning occurs through experiences in communities of practice and is a process of becoming, wherein people become better able to participate in the activities of their community.

Situated learning theory understands learning as a social act where meaning is co-constructed within a community of practice and is contextually dependent. Through social interaction knowledge and practice are maintained and transformed in an ongoing way. In this view, learning is inextricable from practice, and knowledge and action are dialectically related and co-constituting. Lave and Wenger trace how people move into a community of practice through immersion as a newcomer and move through a process of increasing centrality and mastery (Lave and Wenger, 1991). Through active social engagement, newcomers learn the practices and implicitly begin to understand the logic and theory behind the practices and the ways they are organized.

Legitimate peripheral participation describes the process by which new members become masters at activities within a community of practice. Rather than learning through mimicry or through instruction, Lave and Wenger suggest that learning occurs through "centripetal participation in the learning curriculum of the ambient community" (1991, p. 100). Members achieve full participation not only by learning skills or

reproducing types of talk, but also by developing the ability to participate in the full practices of the community. As participants become more acculturated and more competent, they move toward more roles of mastery in the community. Brown and Duguid describe situated learning as "essentially becoming an 'insider'" (1991, p. 48), where new members learn to function in the community of practice.

Communities of practice are understood as "a set of relations among persons, activity, and the world over time and in relation with other tangential and overlapping communities of practice" (Lave and Wenger, 1991, p. 98). These groups of people are mutually engaged in joint work using a shared repertoire of practice (Wenger, 1998). Formally or informally structured, these communities centre on the practice of their particular types of work. Communities of practice co-develop and maintain specific performances of culture and practice. Legitimate peripheral participation is a process of learning the skills and the underpinning logic as novices move into full participation in the community. Wenger claims that "the curriculum is the community itself" (1998, p. 100), thus part of the work is in sustaining the community of practice and building participants' identification within it.

Identity is key within a community of practice, and is understood not as something acquired, but as a process of becoming. Lave (1996) claims that the most significant thing people learn in a community of practice is how to become a full-functioning member. She states that in the process of learning, "people are becoming kinds of persons" (Lave, 1996, p. 157). The focus on learning reflected in identities emphasizes the dynamic ways that identities are produced within certain contexts and communities. Lave and Wenger argue that while participation is critical to engagement and learning in the community of practice, ultimately a "deeper sense of the value of participation to the community and the learner lies in *becoming* part of the community" (1991, p. 111, emphasis original). They claim that newcomers' increasing sense of identity as masters is some of the most significant learning in a community of practice, and indeed, a key indicator that they have achieved centrality.

Situated learning and legitimate peripheral participation are valuable tools for understanding the dynamics within social movement organizations. Few have applied these learning frameworks to social movement contexts, but those who have, including Kirshner (2008), Ebby-Rosin (2005), Evans (2009), and Curnow (2013, 2014) have shed valuable light on the learning dynamics that enable participants to create change in their communities. Learning is a critical, yet under-developed, concept for social movement theorists to take seriously (Foley, 1999; Hall and Turay, 2006), as it has the potential to unlock key questions about how conceptions of justice are co-constructed, how movements are formed, how participants learn social movement tactics, and how communities change as a result of this learning.

Since these social movement spaces often exist outside of institutional schooling and formal training, situated learning is effective in describing and theorizing the ways that people learn and become particular types of activists. Moving toward full participation in an activist community of practice is often a process of learning the particular practices and frames of a social movement organization and moving from peripheral involvement to full participation through active engagement in campaigns. Co-negotiating processes like facilitating meetings, planning, protesting, and coordinating public messaging is a process of participation, and ultimately of becoming – as social movement theorist Charles Tilly argued, a social movement is "what it does as much as why it does it" (in Munro, 2005, p. 75). Situated learning theory allows us to understand how the practices and tactics – that is, what activists do – produce the community, and thus the movement. Therefore, sociocultural learning theories' articulations of learning-as-becoming are particularly apt for social movements.

However, few have looked thoroughly at the ways that activist identities and practices are co-produced and are an ongoing accomplishment within the community of practice. Through this analysis, I show how individuals are brought into a process of legitimate peripheral participation, how their participation produces learning, and how that learning enables new members to become activists and shape the meaning of activism within their community.

Context and Methodological Approach: United Students for Fair Trade

This paper is an analysis of the work of United Students for Fair Trade (USFT). USFT documents define the organization as "national network of student organizations advocating around Fair Trade principles, products, and policies" (USFT 2011). USFT emerged from work to mobilize students around Fair Trade in order to change the purchasing policies of their high schools, community colleges, and universities. Between 2003 and 2006, activists affiliated with USFT ran an estimated 350 campaigns on their campuses. One participant, Katrina, described it saying "USFT was a catalyst for Fair Trade. Our Convergence brought together so many students from so many schools and was the launching point for Fair Trade campaigns. It raised awareness for Fair Trade and brought Fair Trade into the limelight." The student organization played an important role in building demand for Fair Trade Certified products, mobilizing volunteer actions that formed the base of the social movement, and applying pressure to certifying agencies and businesses (Wilson & Curnow, 2013).

USFT developed as a community of practice through their joint work, shared repertoire, and mutual engagement to promote Fair Trade Certified products on campuses across the US. Each of the activists on the Coordinating Committee was engaged both locally in campaigning at their school and nationally as organizers

of students across the country, coordinators of the grassroots campaigns, and negotiators with the other Fair Trade organizations, certifiers, and businesses. USFT was embedded within larger communities of practice as well, including the Fair Trade movement internationally.

An elected Coordinating Committee of students led USFT. The Coordinating Committee was made up of 15-22 students representing different regions of the US and coordinating different core campaigns within the organization. Decisions were made through a consensus-process and the Coordinating Committee was officially non-hierarchical (USFT, 2011). Every summer a new group of students was brought into USFT's Coordinating Committee and transitioned from campus organizing to coordinating internationally. Their learning process through engagement in USFT forms the foundational context for this study.

Data Collection and Analysis

From 2003 until 2008 I was involved in the USFT community of practice. For years I worked as a student activist and later as a professional student organizer with a variety of organizations geared toward ethical consumerism and corporate accountability in the United States and internationally. This research is from an ongoing research project with Dr. Bradley Wilson that examines student activist tactics within ethical labelling movements, the ways that student labour is leveraged and commodified, and the ways student activists learn and act in solidarity with peasant workers and cooperative organizations.

I focused on the USFT Coordinating Committee as a community of practice. Data was collected in 2011 and participants were all former Coordinating Committee members or active affiliates who had been involved from one to five years but were no longer involved in USFT and had not been for four to six years. Twelve participants responded to an extensive, qualitative survey that provided insight into their experiences and development over time. An additional thirteen participants were interviewed in semi-structured and unstructured interviews that lasted one to three hours. Interviews asked detailed questions about participants' involvement, their learning incidents, and their identities as activists. I also conducted a textual analysis of primary documents from the period, including emails and organizational materials, examining the ways the organizational structure, campaigns, and values were described. I used ethnographic notes from the time as well as my participant experiences from the five years I was involved to complement the interview and survey data. I looked for themes in the ways learning, participation, and organizing were formally discussed and codified, as well as the ways they were performed and critiqued.

Responses were coded and analyzed, specifically around issues of activist identity, how people learned practices, perspective transformations, and how the community fostered learning in order to answer questions of how and what young adult activists learned and how their subject-position impacted their learning of skills, identities, and political analysis. Learning was defined as shifts in participation, so when respondents reported changes in their participation or their peers' participation, it was coded as learning. Similarly, identity was coded when respondents mentioned affiliating or disaffiliating with certain labels, communities or practices. Based on repeated review of the audio and transcripts and iterative cycles of coding based on the questions above, I refined my focus, shifting my attention to the ways participants learned specific activist practices and the ways that the core values of the community were revealed through the process. Ethnographic and interview data was anonymized and organized around activist identity development to illustrate the shifts in identity and practice over time as participants moved toward full participation as learners, organizers, and leaders within the social movement organization. The findings were triangulated across multiple interviewees and are consistent with either the texts, notes, or my participation, or a combination. Once an initial analysis was completed, the draft was circulated among participants, five of whom responded with written or verbal feedback, and their critiques of the data and my analysis have been integrated into the results that follow.

Doing Learning, Doing Activism: Techniques, Norms, Rituals

Each year, a new cohort of student leaders were brought into USFT's Coordinating Committee. They reported facing a steep learning curve as they moved from organizing on their own campuses to nationwide organizing and coordinating at an international level. Respondents reported learning many things, including about Fair Trade, international development, and the coffee industry. They also reported learning how to manage non-profit organizations, fundraise, plan major international events, and negotiate with large corporate actors at high levels. In the sections that follow, I will show how burgeoning young adult activists took gradual steps as they learned increasingly more about the activities that were expected of them, and through their increased activity, how their consciousness developed. New coordinators and returning coordinators learned through their activity, co-constructing the organization and co-creating their identity as Fair Trade activists.

Group Agreements

At the most basic level, newcomers were introduced to techniques, norms, and rituals that facilitated group process and built a sense of shared values in action. Isadora described how in almost any space a newcomer entered, the group process would begin with a discussion of "ground rules". Though they were collectively

generated by asking participants for suggestions of rules which were agreed upon with each new group, these ground rules generally consisted of the same core agreements. With remarkable frequency, these same basic rules were offered and agreed to in almost any group meeting, workshop, or conference. Typically, someone would offer a version of "listen to understand". Someone else might add a statement of confidentiality or not using people's names when sharing conversations outside of the group. Another typical offering was "step up, step back", a request that people who may be more assertive or more talkative in a group step back. These ground rules were emergent in nearly every workshop space and meeting, yet reflected the core practices and principles of the community of practice, as well as the relationship between practices and principles.

The ground rule listen to understand, for example, was not just a rule about paying attention, but rather underscored an organizational emphasis on consensus-building based in relationships. In interviews, members articulated that this rule was established in opposition to what was perceived as a common practice of waiting for one's turn to speak and argue rather than really attending to what a speaker was trying to communicate. However, this logic was rarely made explicit in ground rules. Members were expected to comply, and presumably learn the value of the approach through their participation.

Similarly, step up, step back worked to shift the practices of the group so that they reflected the equity commitments of the organization. Isadora claimed that the goal was to create space for less dominant or less vocal people to step up, and to encourage less vocal people to take a risk, as it were, and contribute orally to the group process. Isadora identified that this practice was underpinned by an anti-oppression logic — that White privilege and male privilege were enacted through dominance in participation, and that intervening through an explicit guideline was meant to make people think about "how much space they took up" and to subvert what might otherwise seem like a natural process. Miles also commented that the expression of privilege was being addressed through these types of rules without being labelled as such, where the goal was to make it more likely that women and people of color would be more active in conversations. Though this was not made explicit, people were expected to conform to the ground rule and understand it through their practice.

The co-construction of group agreements served multiple purposes. The first was governing group behaviour. Through the process, newcomers could see in normal interactions what the expectations for behaviour were, could practice operating within those norms, and eventually begin to offer the standard rules if they had not yet been incorporated into the group agreement as a way of contributing and potentially innovating. Participating in the co-development of group agreements was a basic way people could demonstrate their proficiency in the cultural production of USFT in small ways, regardless of their level of engagement. The second purpose of ground rule setting was in modeling facilitation strategies. Through this mechanism, newcomers could observe the facilitation techniques of more central members and participate in meetings structured around the establishment of the shared agreements. I observed that newcomers were active participants in constructing a space that was facilitated and enabling the facilitation. Additionally, they could see explicitly what should be enforced in meetings and workshops, and how members could be corrected or disciplined. Randy reported that through this process, newcomers were also able to learn, through their participation, which norms were stated but rarely enforced by watching others breach or by breaching themselves. From my experience, this breaching and repair process was most likely regarding step up and step back, where if someone was perceived to be taking up too much space, the group might be reminded of the rule generally by a facilitator or participant, a facilitator might intentionally avoid calling on that person, or the individual might be told directly to step back. The participation and observation in the group agreement process enabled newcomers to eventually take on the facilitation of sections of meetings, and later entire meetings, workshops, or conferences.

Keeping Stack

In much the same way, newcomers learned from more experienced peers about other processes that were central to the organization's activist culture. One way activists managed contentious conversations was through a technique they referred to as "keeping stack." Dani described how she learned stack, saying,

> Darren, another student leader, had been involved in more radical activism in his hometown, so when we were having a hard time managing some difficult discussion, he introduced the idea of using stack. He explained how one person would keep stack, basically keeping track of who wanted to respond to a specific idea, who wanted to speak next. And as part of it, people would wiggle their fingers or snap if they agreed with what someone was saying, so that then they wouldn't have to interrupt the flow of the discussion. So Darren kept stack for us, and by watching how he did it, we learned how it worked and how to keep stack. And once we got the hang of it we used it a lot and took it back to the rest of the organization.

Dani described a process in which a new skill was introduced; novice coordinators learned through instruction and modeling how to use the skill. This included experimentation with the hand signals and the timing of the

complex system—to indicate they wanted to speak, a participant would raise their hand or index finger and wait to be acknowledged by the person keeping stack. Once acknowledged, the person would wait until their name was called in the order the facilitator collected the names, unless one had a "direct response", indicated by alternately pointing both index fingers at the speaker. Direct responses did not join the queue, but instead were able to speak immediately following the speaker they sought to respond to. Additionally, if someone agreed with what a speaker was saying, rather than joining the queue to voice their support or responding directly, they instead raised their hands and wiggled their fingers. There were also signs for interrupting for a process question, indicating confusion on a topic, raising volume, and other requests. These signals could be difficult for new members to deploy at the right time and difficult to keep track of the full inventory of options one could use for a given situation, but in order to participate, new and experienced members needed to use signals at the appropriate time, especially if they wanted to speak by joining the stack. Based on my observation, as newcomers became more proficient at participating using this process, they had opportunities to try to keep stack, thus learning more about how it was done in practice, but also shaping the way others in the collective understood how stack was kept. Becoming the stack facilitator required greater proficiency at knowing what certain signals indicated and when they should be used, but also gave the facilitator the freedom to make accommodations based on their own facilitation practices and the participants' usage of signals. Later generations would not learn through explicit instruction, but would simply be immersed in the practice. They would learn through observation and experimentation with occasional coaching or correction by the facilitator or their peers.

The process became an important performance, a skill that new coordinators needed to be able to use appropriately when it was being deployed. Members had to know how to indicate they wanted to speak in order to get into the stack, so in order to participate in discussions they had to be proficient in the performance of stacking. Members also had to respect the stack order; interrupting the order to express an opinion, which would be acceptable behaviour in many other argumentation contexts, would be considered rude and entitled, so one needed to adhere to the process. It was also a skill that a novice would be expected to employ as a facilitator as she or he moved toward full participation, which meant learning how to identify speakers jockeying for position in the stack and managing direct responses to previous speakers.

Conference Calls

Conference calls were sites of key practices to perform, and several interview respondents joked that this was the skill that was most important to learn. Most of USFT's decision-making was conducted via conference calls, since coordinators lived across the United States. Having 15-20 people participate on a conference call required a lot of shared understanding around process and facilitation, according to Isadora. Throughout my data, people suggested that almost none of this was taught in an explicit way. Rather, when a new generation of Coordinating Committee members were on their first call, an experienced facilitator would conduct the call. He or she would likely begin by explaining basic call etiquette, like introducing oneself before speaking, until everyone got to know everyone else's voices, or announcing oneself when you joined a call, or muting the phone while you are listening. Beyond that, though, people had to learn by listening and participating in calls. The next call might also be facilitated by a more experienced facilitator, but quickly new coordinators were expected to begin facilitating calls, as the responsibility rotated through the group. Novices learned that they should email out a request for agenda items by receiving requests via email. They also learned that facilitators were expected to develop agendas that included specific items with allotted times to each item by receiving proposed agendas. Rita noted how people had to learn how to participate, and described her own process of development. At first, she would only volunteer to take notes. She said, "I felt a sense of responsibility, this is my contribution to keeping the ball rolling, making the facilitator more successful… that's how I learned about facilitation." She very clearly identifies how her peripheral participation as a note-taker enabled her to learn the central skill of facilitation. Novices also learned specific ways of facilitating, managing time, and calling for and conducting a consensus process for decision-making. What was not taken up through imitation might be assisted or corrected by one of the old-timers of the community of practice. Occasionally, too, if someone was deemed to be doing too bad a job at conducting a call, other people would usurp the responsibility, intentionally demonstrating, either implicitly or explicitly, the "right way" to facilitate. These lessons of what was correct protocol were supplemented by the many online instant messages that some coordinators, including Katrina and Roxana, reported using while simultaneously on conference calls. Commiserating about bad facilitation or making fun of awkwardness reinforced the norms around what was good and bad facilitation. Throughout the data, participants stressed that becoming a strong facilitator was one of the ways that coordinators established themselves within the Coordinating Committee and became recognized as student activists.

Ideological Underpinnings

These practices governed group dynamics. Emergent activists learned these skills through their participation in the group. By watching and participating in the processes as participants and as facilitators, they became

activists. As they adopted these frames of understanding, these techniques, and these rituals, they became more fully able to function as full activists, not only in USFT, but also in the broader North American leftist activist culture. As young adult activists learned these skills, they moved from peripherality toward full participation within USFT. Through legitimate peripheral participation, new members were able to observe the practices of a community and participate in the activities. As they became more central, they would facilitate small pieces, and then be expected to facilitate with support, and then be able to facilitate independently. This "leadership ladder" as USFT referred to it, was a process of immersive skill building and acculturation that enabled people to do the work proficiently while coming to understand the political theory of the organization.

Through engagement in all of these processes, the underpinning logic and theory of social change for USFT could be revealed. The coordinating ideology was about producing facilitators attentive to interaction and the production of anti-oppressive spaces in order to enable radical democracy, consensus-based decision-making, and non-hierarchical participation. Lee said that USFT's facilitation processes:

> Focused on an obsession with empowerment. We were talking about empowering producers of coffee in the Third World. And if we were going to live that ethic with integrity, that meant remaking our own selves and the way we interacted in a way that was more empowering of everybody. We're talking, like, making sure voices that are traditionally marginalized are really heard, at every level. That meant a lot of time spent organizing minority caucuses in USFT, making sure that less extroverted people on conference calls were heard.

Through developing ground rules, the community of practice developed a political critique of privilege and worked to ensure that all participants could be involved in an inclusive community in order to foster truly democratic decision-making. Keeping stack served a similar purpose, and worked to enable a consensus model of decision-making. Coordinating conference calls the USFT way ensured that participants had the information they needed in order to make decisions well, while also attending to the ways that privilege was enacted during the meetings. Randy said, "USFT was more concerned with how they did things than other activist groups I had been part of, which were mostly results-oriented without a whole lot of concern about their method… It was striking how much people were concerned with attitudes and the way things were said or done." All of USFT's facilitation strategies were developed from a political ideology and worked to engrain that ideology in the community members. Newcomers' active participation enabled particular forms of learning, and it was the demonstration of both the performances and the political underpinnings that allowed people to claim roles of centrality within the community of practice.

One reason that these performances of facilitation and group dynamics were so important was because they simultaneously signalled and constructed the broader politics of the group. USFT had a stated commitment to building an anti-oppressive movement, internally and externally. Their external politics, like campaigns, gave them a public facing approach to anti-oppression, but both internal and externally oriented practices had to reflect their core values, through what is referred to as prefigurative action (Breines, 1989), which gave them an immediate outlet to enact their values. The group dynamics and processes underlined here demonstrate the ways that people were eased into these politics in an internal community of practice as a way of developing both the skills they would need to coordinate the public facing campaigns, and the political/ideological approach that USFT believed should underpin the external work. Within USFT's community of practice, demonstrating proficiency in the internal process of legitimate peripheral participation unlocked people's ability to become full participants as activists in the external campaigns.

On Becoming an Activist

In my interviews, I found high levels of discomfort with the idea of activism when participants described their entrée into USFT. Most respondents stated that when they became involved with the organization, they did not identify as activists. This is particularly significant because all of these individuals would eventually take on leadership positions in the organization. Rita said, "I didn't like the connotation of the word activism, I did not consider myself an activist. It's kind of funny I was spending eighty hours a week, like every amount of free time I had, living, breathing and thinking about social issues and how to change them." For Coordinating Committee members, their initial lack of identification with the label of activism had to do with preconceived notions of activism. Katrina said, "Working with USFT helped me get past my stigma of activism and feeling radical enough" indicating her dis-identification with the label. These new members shared a commitment to the cause and were deeply engaged in the work, but initially did not understand the work as activism. Lizzie said "It took me a while to recognize that I was an activist… enough people started telling me I was an activist. I was like, I'm just doing stuff I like." Lizzie and others identified with the work first, seeing its relevance to their lives and believing it created a real impact in the world. For Lizzie, when other people from inside the community identified her as an activist leader, she began to see herself that way as well. On a similar note, Rita later said

> It was more about participating in the community. I didn't think the things that I was doing were activism, it was the task I was doing. I just didn't think I was an activist, I wasn't averse to it. It changed after the Convergence – I felt more embedded in the community, like this is an activist cause I would want to be involved in regardless of tasks… It grew on me.

She focuses on the ways that the work and the community provoked her to change her ideas about activism and see herself as a contributing member of the activist community. For these members, it was only through their engagement in the community doing the work they felt was necessary for a cause that they felt affinity for that they came to identify themselves as activists.

All respondents stated that after their involvement on the Coordinating Committee, they did self-identify as activists. For many, immersion in USFT changed their conceptions of activism, grounding them in real-life activity and de-mystifying the idea of activism as fringe or fanatical. Roxana said:

> It made me realize that activism is a lot more than just protesting or deciding one day to sit in the front of the bus. It's about movement building and sustained effort in order to create any real change. Activism is also about educating and empowering others in order to build people power and, ideally, create systemic change in the long run.

She notes a significant change in her consciousness when it came to the idea of activism – from the activity to the strategy to the broader work of consciousness-raising. Beyond the facilitation work, USFT's organizing practice, campaign strategies, and education pedagogies were also highly specified, constructed, and maintained through the community of practice, but an analysis of these is outside the scope of this work. Future work will entail an analysis of the ways that USFT's campaigning and pedagogy works to produce activist identities.

USFT student activists carved out a particular identity as activists, one that was specific to Fair Trade. Their approach to facilitation was highly specified, as was the theory of change and anti-oppression that was the foundation or their work. Through their work, participants became able to identify themselves as activists because 'activism' now had a particular meaning for them that was rooted in their tasks and their engagement in the community. Rather than an abstract term connected to extreme political expressions, 'activism' became a concrete performance of facilitation and social change work that was rational and that they were capable of doing. Their identification with the community of practice drove them to re-evaluate their earlier opinions of activists. Within this space, the process of becoming an activist constituted taking up the techniques, norms and rituals of the community. Through their process of adopting the practices, participants learned the logic behind the practices and began to identify themselves as activists.

At the same time, though, the specific USFT activist identity was rooted in larger communities of practice; the forms of participation that newly labelled activists became skilful at performing were signifiers within student activist communities across North America. Many of the practices analyzed above were not exclusive to USFT, but spanned other youth social movements, including the alter-globalization and anti-sweatshop movements. Part of what drove people's identification as activists, then, was also their ability to situate themselves within a broader community of practice. When USFTers who previously had not thought of themselves as activists took on these practices, they could see how they were, in fact, acting like activists. By removing the stigma that respondents identified and focusing on manageable tasks, participants gradually worked toward proficiency and centrality that slowly made it possible for them to see themselves as others saw them.

This case has demonstrated a process of situated learning, wherein students involved with USFT moved from dis-identification with the label of activist to a strong identification with the term through their shared work in a community of practice. Through legitimate peripheral participation in the basic facilitation strategies of the organization and their ongoing work in the community of practice, these activists shifted their identities and their abilities to participate in social change work. Through immersion and experimentation with the practices of the community, including developing group agreements, keeping stack, and participating in and facilitating conference calls, new members of USFT Coordinators' community of practice became activists. Their performance of these practices as peripheral members allowed them to learn and shape the ideologies that were foundational to the practices. Through that learning and members increasing ability to perform the full practices of the Coordinating Committee, they became recognized as full members and understood themselves as student activists.

Situated learning frameworks have much to offer social movement theorists and activists who want to understand the ways that new members become embedded in social movement communities of practice. This case shows the value of situated learning theory for the study of learning in social movements, in which people often learn through loosely structured engagement in a committed group of volunteers and learn the practices of the community through their shared work as they attempt to change the world. Legitimate peripheral

participation describes the learning trajectories in this social movement organization, as new members gradually became more fully immersed in the practices and better able to perform them. Participants' identity development emerged from their experiences of participation and co-construction of the community and its practices. Situate learning opens new avenues for social movement researchers to understand why people join movements, how movements evolve, and how frames and ideologies are constructed and circulated by activists in their daily activity.

For the Learning Sciences, this case offers a new context to understand learning and identity development, and links social movements with situated learning. This bridge strengthens our ability to understand learning in formal and informal learning contexts, and offers new sites of inquiry for researchers interested in the relationships between communities of practice and social change. Additionally, this case draws attention to the political nature of communities of practice and how ideologies are developed, propagated, and maintained in communities of practice. USFT's case also helps us to theorize how new members of communities develop identities through their participation, and explores how practices are passed on over generations within a community of practice. Finally, using Learning Sciences approaches in the context of social movements allows us to see the impacts movements have on individual activists engaged in collective action and to understand how social change is produced through participation in communities of practice.

References

Breines, W. (1989). *Community and organization in the New Left, 1962-1968: the great refusal*. Rutgers

Brown, J.S. and Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning and innovation. *Organization Science.* 2:1 40-57.

Curnow, J. (2013). Fight the power: situated learning and conscientisation in a gendered community of practice. *Gender and Education*, *25*(7), 834-850.

Curnow, J. (2014). Climbing the Leadership Ladder: Legitimate Peripheral Participation in Student Movements. *Interface, 6*:2

Ebby-Rosin, R. (2005). *New voices at the table: Organizing low-income Latino parents for school reform* (Unpublished doctoral dissertation.). Stanford University, California.

Evans, M. P. (2009). *Inside education organizing: Learning to work for educational change* (Unpublished doctoral dissertation.). Boston College, Massachusetts.

Foley, G. (1999). *Learning in Social Action: A Contribution to Understanding Informal Education. Global Perspectives on Adult Education and Training*. St. Martin's Press, Inc. New York, NY

Hall, B., & Turay, T. (2006). State of the field report: Social movement learning. *Victoria, BC: University of Victoria*.

Kirshner, B. (2008). Guided Participation in Three Youth Activism Organizations: Facilitation, Apprenticeship, and Joint Work. *Journal of the Learning Sciences*, *17*:1, 60-101.

Lave, J. (1996). Teaching, as Learning, in Practice. *Mind, Culture, and Activity, 3:*3, 149-164.

Lave, J. and Wenger, E. (1991). *Situated Learning Legitimate Peripheral Participation*. NY, NY: Cambridge.

Munro, L. (2005). Strategies, action repertoires and DIY activism in the animal rights movement. *Social Movement Studies, 4*(1), 75-94.

United Students for Fair Trade. (2011) www.usft.org/about. Accessed on April 1, 2011.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity.* Cambridge: Cambridge Press.

Wilson, B. R., & Curnow, J. (2013). Solidarity™ : student activism, affective labor, and the fair trade campaign in the United States. *Antipode*, *45*(3), 565-583.

Acknowledgments

Thanks to USFT and Bradley Wilson for their collaboration on this project, as well as the GLITTER Lab, Kate Curnow, and Andrew Kohan for their comments on earlier drafts. Thanks to all of the respondents for their participation.

"We Should all Help Each Other": Latina Undergraduates' Practices and Identities in the Figured World of Computing

Heather Thiry, University of Colorado, Boulder, UCB 580, Boulder, CO, 80309, heather.thiry@colorado.edu
Sarah Hug, University of Colorado, Boulder, UCB UCB 322, Boulder CO, 80309, hug@colorado.edu

Abstract: The number of Latinas earning computing degrees and entering technical careers is stubbornly low. This study uses Holland, Lachicotte, Skinner, and Cain's (1998) concepts of identity and figured worlds to explore the experiences of 22 Latina undergraduates in computing majors. Using semi-structured, focus group interviews, this paper describes participants' identity production as empowered computer scientists. Results indicate that Latinas faced many cultural constraints within the landscape of computing, including isolation, marginalization and microaggressions, yet they also described practices and relationships that helped them to persist in their majors. Successful disciplinary performances and access to Latina role models were pivotal in students' adoption of empowered identities. Study participants challenged the notion of computing as a competitive, individualistic enterprise that permeated the local and global computing communities in which they operated. Instead, they developed identities as engaged, community-oriented computer scientists and enacted these identities through their everyday practices in their departments and in the local community.

Introduction

Learning involves more than the acquisition of technical knowledge and disciplinary expertise. Though these elements are important aspects of the learning process, learning also involves identification with a discipline, profession, or community; in short, "becoming" a certain type of person. In scientific, mathematical, engineering, and mathematics (STEM) disciplines in higher education, learning encompasses an interrelated process of developing accountable disciplinary knowledge, identifying with the discipline, and navigating institutional and disciplinary contexts (Stevens et al., 2008). While undergraduate students must successfully navigate institutional and disciplinary benchmarks, such as core courses and requirements, the process of *becoming* a scientist occurs within the everyday practices, relationships, and interactions that students experience in the cultural context of their discipline.

This paper focuses on the underrepresentation of Latinas within the field of computing and explores the ways in which Latina undergraduates come to see themselves, and to be seen, as successful and empowered computer science students. The larger purpose of the study was to understand how Latinas negotiated disciplinary cultures that subordinated them as women and as individuals of Hispanic origin. Through interviews with Latina computing majors, we explored the barriers and supports they encountered in authoring competent identities in technical fields. The research questions that drove this investigation were:

- How do Latinas negotiate the figured world of academic computing, particularly in regard to their computing expertise?
- What social and cultural practices support Latinas in taking up professional identities in computing?
- How does the intersectionality of race, class, and gender influence the position of Latinas in the figured world of academic world of computing and the identities that they craft within the discipline?

Social and Cultural Barriers in STEM disciplines Faced by Women and Latino/as

Pathways in STEM fields are especially difficult for underrepresented minority students and women (Gasbarra & Johnson, 2008; Wajcman, 2010). Women are underrepresented in scientific, technology, engineering, and mathematics (STEM) careers in 121 developed and developing countries (UNESCO, 2009), within the European Union (European Commission, 2009), and in the US (National Science Board, 2012). Inequities in computing—the disciplinary site of this study—are especially grim. Only 7% of baccalaureates and less than 1% of doctorates in computer science in 2012 were granted to Hispanic US citizens (National Center for Education Statistics, 2012), though Hispanics represent 16% of the total US population, and nearly one quarter of the youth population (US Census, 2011). The National Center for Women and Information Technology (NCWIT) reported that only one percent of the information technology jobs in the US are held by women of Hispanic origin (NCWIT, 2004).

Historical-cultural views on what kind of work is appropriate for women foster gendered expectations that, in turn, profoundly affect the numbers of women in computing fields. For example, in the US, the dominant view of computing is that it is an asocial, highly technical, masculine pursuit (Barker & Aspray, 2006). A

number of processes including early socialization (Clewell & Braddock, 2000), perceptions of and actual instructor bias (Beyer, Reynes and Haller 2004), and negative departmental and classroom climates (Margolis & Fisher, 2002; Seymour & Hewitt, 1997), work in tandem with these historically and culturally engrained views to discourage women from pursuing computing degrees and careers.

In addition, Latino/as face a number of sociocultural, economic, and educational obstacles in higher education degree attainment (Hurtado, Carter, & Spuler, 1996; Oseguera, Locks, & Vega, 2009). Lack of faculty support, discomfort on the university campus, and financial struggles contribute to high attrition rates (Gloria et al. 2005; Santiago & Treindl, 2009). In STEM disciplines, cultural incongruence between minority communities and academic departments has been argued to contribute to the underrepresentation of Latino/as, African-Americans, and Native Americans in these majors (Bonous-Hammarth, 2000; Cole & Espinoza 2008). Latino/a students are less likely to have completed college-ready coursework upon high school graduation (Tyson, et al., 2007). Additionally, K-12 schools with sizeable numbers of students from non-dominant communities emphasize basic skills over higher-order thinking (Sleeter, 2005). Racial and ethnic disparities in computing education can also be attributed to the stubborn persistence of the digital divide (Margolis, 2008; Warschauer, Knobel & Stone, 2004).

Research on STEM education has elucidated historical and cultural factors that contribute to the lack of women and minorities in technical fields and identified strategies for supporting students once they have enrolled in STEM majors. However, researchers have not often focused on the ways in which specific underserved populations in STEM, such as Latinas in computing, experience intersectionality, negotiate barriers, and craft identities in fields in which they are overwhelmingly underrepresented. In this paper, we identify the life history experiences that are associated with the development of empowered identities for Latinas in computing fields.

Conceptual Framework

We use Holland, Lachicotte, Skinner, and Cain's (1998) concepts of identity and figured worlds to explore the experiences of Latinas within computing majors. Identity is formed within certain figured worlds through routine interactions, activities, and relationships. Figured worlds are socially and historically situated realms of human activity with their own sets of values, norms, and expectations. Figured worlds may be broad, such as academia, or local, such as a student club within a campus department. People enter into or are recruited into figured worlds and they "come to identify themselves as actors of more or less influence, more or less privilege, and more or less power in these worlds" (Holland et al, 1998, p.60).

Identity is a valuable construct for interpreting the experiences of actors with less privilege and power within dominant cultural contexts, such as Latinas in computing. According to Holland, et al. (1998), people may challenge taken-for granted notions and cultural constraints through the process of identity development. For example, Latinas may accept traditional, masculine notions of a computer scientist as a white or Asian male "hacker" or they may negotiate new definitions of computer scientist as activist and engaged with youth, education, or community outreach. Identities, including ways of becoming a computer scientist, are not static and fixed, nor do they spring from some "essential" characteristic such as gender or race. Instead, identities are works in progress and are formed in and through everyday practices and interactions.

Identity involves not only actively identifying oneself as an actor within a specific social and cultural field, but also being seen by others as such. While this notion of identity allows for agency within peoples' lives, individuals are not free to adopt any identity they want. Indeed, they are constrained by given social and historical conditions; but, they may also "improvise" and act creatively within those conditions. The concept of identity within figured worlds allows us to examine hierarchy, status and power within specific cultural realms of activity, such as academic departments or disciplines. This lens provides a framework for exploring the intersection of agency and structure and the ways in which individuals negotiate power within cultural domains. In this study we use the concepts of identity and figured worlds to explore how Latinas navigate oftentimes difficult academic pathways and craft identities that value the intersectionality (Crenshaw, 1991) of being Latina in the figured world of computing.

Research Design and Data Sources

We conducted semi-structured interviews to investigate—from the perspective of Latina actors in the figured world of computing—the interactions, practices and relationships that facilitated or hindered their identification with the discipline. All of our study participants were involved in the Computing Alliance of Hispanic-Serving Institutions (CAHSI), a National Science Foundation-sponsored consortium of ten Hispanic-serving Institutions (HSIs). CAHSI implements a number of pedagogical innovations that support the recruitment, retention, and advancement of Hispanics in computing. Participating institutions span the cultural and regional diversity of Latino/as in the US, from a university in Puerto Rico, to ethnically diverse urban institutions in California, to

universities in Florida with significant numbers of Cuban, Central, and South American students, and border universities in Texas and New Mexico with large populations of Mexican-American students.

Latina undergraduates are severely underrepresented in the fields of computer science and computer engineering. For example, Latinas accounted for less than one percent (only 64 out of 8977 total graduates) of bachelor's degrees in computer science or computer engineering from U.S. institutions in 2012 (NCES, 2012). CAHSI serves as a national hub for Hispanic computing students, faculty, and professionals to support and network with one another. Due to the dire underrepresentation of Latina undergraduates in computing fields, CAHSI was an ideal site through which to identify research participants.

We conducted focus group interviews with all of the female participants from CAHSI (n=22) at an academic conference for underrepresented minority students and faculty in STEM disciplines. We sought to interview all female participants at this professional meeting to not only explore Latina's experiences as computing students, but to investigate how their participation in the professional conference may have influenced their identification with the discipline. We scheduled focus groups with all 22 women from seven institutions who attended the conference. Participants met with female peers from their institution when applicable, in groups from one to six total participants. Two interviewers conducted the focus groups, with one researcher chosen beforehand to lead the discussion. Of the 22 female focus group participants, almost all were undergraduate students (81%) and self-identified as Hispanic (77%). Almost all participants were enrolled in computer science (CS) or computer engineering (CS) majors, although two students were enrolled in computer information systems (CIS) and one was a mathematics major with an emphasis in computer science. Interviews lasted 45 to 90 minutes, were digitally recorded and transcribed verbatim. Interviews focused on participant's experiences within their departments and the larger field of computing. Sample interview questions include: "How did you first become interested in computer science?" "Where do you 'fit' in your department?" "Do you see any barriers that impact you as a Hispanic woman in computing?" and "What has supported you in your pursuit of a computing degree?" All research procedures were approved by the human subjects institutional review board. Pseudonyms have been used to protect the confidentiality of participants.

Analysis Methods

Interview transcripts were coded using domain analysis (Spradley, 1980). Researchers searched for units of meaning within the data, coding interview transcripts for examples of "cover terms" within broader "domains." Taxonomies were then constructed linking coded examples to domain categories through a semantic relationship such as "is a kind of" or "is a way of doing." Domains were generated both deductively, based on our research questions and our conceptual framework, and inductively, based on emergent themes from the data. For example, some deductive domain categories from our analysis include: *Barriers in Computing, Supports in Computing, and Professional Identity.* Some inductive domain categories include: *Gendered Expectations* and *Family Influence.* Two researchers generated the initial codebook in *NVivo* qualitative software based on the research questions and conceptual framework. The researchers initially coded transcripts in tandem to identify emerging domains and to gauge inter-rater reliability. The researchers then divided the rest of the interviews and coded them separately, eventually merging them into one *NVivo* project. Throughout the coding process, the researchers met regularly to discuss emerging findings.

Results

Power and Status in the Figured World of Computing

Through the formal support provided by CAHSI and informal support fostered within peer networks, some women experienced shifts in identity similar to the Chicano activists/educators described by Urrieta (2007), who took up an activist identity with a desire to educate others and give back to their community. However, the life experiences that fostered these shifts among our study participants were different from those in Urieta's research, with the exception of the experience of oppression.

As Latinas in computing, the undergraduates described their experiences of oppression in their discipline. They discussed isolation, marginalization, and microagressions, meaning "brief and commonplace daily verbal, behavioral, or environmental indignities" that highlight difference and power (Sue, et al., 2007 p.271). The women in this study encountered these slights in both local and global figured worlds of computing and often narrated their experiences with a focus on gender. For instance, Elisa recounted an experience at a conference where she felt ignored and invisible to a male professor, in stark contrast to the attention he gave to her male peers. This microagression highlighted her subordinate status as a woman in the realm of computing.

> When we come to other conferences or other universities, I sometimes feel like, "I'm a woman." I've run into people, like other professors at other universities, and they don't pay attention to things I have to say. I just had that experience yesterday, actually. I was really

bummed out. I had to introduce myself to a PhD professor, and I was explaining, "This is what I'm interested in." He went on to the next student, went to the next student, and then he comes back to me and he was like, "Did you already introduce yourself? Did you already tell me about yourself?" He gave everybody information about scholarships, and I was like, "I'm the only one without a paper." I don't know if it was because I was female or what. Just little things… it's not very nice.

The women also became aware of their subordinated status as women in computing through difficult interactions with male peers within the competitive landscape of undergraduate computing classes, as described by Terese and Josefina. Some male peers accepted them as computer scientists, while others held lower expectations or were dismissive of them. These interactions highlighted the devalued status of women within computing.

Josefina: One thing that I notice, in any of my CS classes, there's a few guys that treat us maybe as equals. There's like five guys that are like, "Yes, they can do what we can do." But then the other ones, it seems like you have to do better. You have to study harder and you have to get better, or the same grades to be considered like, "oh, maybe they can do it." Then sometimes it's like, still not, they're still just a girl.

Terese: Most of the time I do better than all of them… But at the same time, it's this disconnect. It's this competitiveness where they don't want anybody being better than them, especially females. But it shouldn't be like that. We should all help each other.

In the exchange, Josefina and Terese lament the lack of support from male peers and contrast the value that they place on community with the competitive identities of many of their male peers. In turn, they propose new ways of becoming a computer scientist within academic spaces.

In these cases, and many others recounted by our focus group participants, gender—rather than ethnicity, class or other socially constructed categories—became the salient aspect of their identity through which they initially experienced power relationships in computing. Their focus on gender subordination may result from their attendance at Hispanic-Serving Institutions (HSI) in which Hispanics comprise at least 25% of the student population. Additionally, they participated in CAHSI, an organization dedicated to creating a community of Hispanic computing students, faculty, and professionals.

Becoming Computer Scientists

Identity shifts that expanded women's notions of what it means to "be" a computer scientist, and "who" can be a computer scientist, were often fostered by interactions with role models. Interactions with role models also led the students to a greater understanding of the intersectionality of their identity within the landscape of computing, and the complex way in which class, race, gender, and language may all play a role in subordination. Participants came to understand the role that race, ethnicity, nationality, and other factors played in conferring diminished status in the field of computing. For instance, Julia commented on how hearing role models' stories at conferences helped her to situate herself within the field and envision that it is possible to be a successful Latina computing professional.

Julia: I think the conference is really good about exposing young females to PhDs that are also females, and have prominent positions in the universities. I think that's really exciting. I come from a university where I'm almost the only girl in all my classes, and I'm particularly the only Hispanic girl, sometimes even the only American. It's nice to see people that are like me, who have similar backgrounds, who have succeeded. It's very inspiring, because you go through all these troubles, and you're starting out with your family, and not having too much money and all that stuff. You hear their stories, and they sound the same. I came from that background. Seeing people who have accomplished that, and who have gotten a PhD, it's very inspiring, to a lot of girls.

Emma: It's really exciting. Sometimes you just need to see that it's possible.

Benita: It gives you hope.

Latina role models helped the women to frame their experiences of oppression in terms of the intersectionality of race, class, and gender. Some of the women also credited the systemic support they received from CAHSI—which facilitated their involvement in disciplinary research, travel to conferences, and interactions with Latina role models—with their consciousness of underrepresentation. Dolores commented that access to the tools and practices of her discipline in her research experience had helped her successfully navigate the academic pathway. Her research mentor also raised her awareness of the underrepresentation of Latinas in computing.

> I love my degree and I'm really glad that CAHSI is helping our communities because we were talking about it in the research lab a few days ago, that not only the fact that there's no women [in computing] but no Hispanics, and [my mentor] was talking about how underrepresented we are. I'm glad that CAHSI is actually helping, because it feels encouraging that we actually can get somewhere, right? Because with no help, I mean, nobody would get anywhere. I'm really thankful for the opportunity.

Thus, many of the Latina undergraduates became critically aware of social and cultural inequities within the figured world of computing. Some women attributed their persistence in the major to this process of identity production, where they became critically conscious of power within the realm of computing, yet also gained confidence that they could successfully navigate—and possibly subvert—those power relationships. Participants gained confidence through successful performance in the field and recognition, similar to the women of color in Carlone and Johnson's (2007) study of science identity development.

As the Latinas in our study displayed competence and mastery in their disciplinary knowledge through research experiences and conferences, not only did they come to see themselves differently, others began to view them differently as well. Rosa described a shift in the way that her male peers perceived the Latinas in her department. The student's engagement in professional activities had conferred status to them within the local figured world of their department and others began to see them as successful computer science students.

> Some of (our male peers), they don't know us, and I guess they don't know what we're doing and, I guess they judge us. Or they have this stereotype of the girls who are not good enough. …they treat me with respect now, *'Oh, wow, you're going to conferences, that's really cool. Wow, I've never done that before. You must be really smart, you work really hard*.'"

In this way, many participants gained status within their departments, seeing themselves, and being seen, as successful computer scientists.

Crafting Empowered Identities

At four institutions, these shifts in identity sparked a more empowered stance among participants. As already noted, the women often felt isolated, disenfranchised, and disrespected by some of their male peers and in some interactions outside of their departments. As a result of their critical awareness of the intersectionality of gender and race within the figured world of computing, they began to subvert these power relationships by creating academically-oriented clubs and other supportive peer networks. Terese, one of the founders of a computing club on her campus, described its goals of creating a stronger sense of community in the department, not just for women, but for all students.

> Basically the goal is to create this sense of community within the department for students to feel comfortable. We have study sessions before the meetings, so students can just come, sit together with other students, and do their homework together. It's great because within that mix there is some lower classmen and upper classmen, and so there's always help there, because everybody is friendly.

Terese reflected on her reasons for attending a conference dedicated to the advancement of underrepresented populations in scientific fields. Her motivations reflect her burgeoning understanding of the complex intersecionality of gender and ethnicity within computing. She hoped to take her new understanding of the positionality of Latinas in computing to encourage underrepresented minority girls in her local area to enter into and persist in STEM fields. Though she embraced some aspects of the traditional culture of computing, and had a professional goal of working at Google as a programmer, Terese had begun to incorporate outreach, mentoring, and local activism into her computer science identity.

> What I can take back [from the conference] is getting girls and minorities [into computing]. Where we work is a HSI, Hispanic-serving institution, primarily, so since not only are women underrepresented but also minorities, what we can take from this conference is to help the students and the girls and the kids in our area. That's what I'm hoping to get out of the conference is new ideas, because it's so hard to keep using the same old ones. When you feel like they're only helping so much.

On another campus, the women negotiated a new organizational identity for their campus computer science club. Aida described how the former student president of the club was a "little iffy, and he was only wanting to do game stuff. He was like, 'We're going to have some LAN parties, and just get together and game all day.' That was all he was about. I was like, 'Yes, that would be fun to do, but we need to do other stuff.'" A group of undergraduate women, mostly Latina, assumed the leadership of the club and created a space where outreach, community service, and professional development were valued. They shifted the values of the club away from an individualistic focus on technical proficiency to a focus on community.

> Last year, and this year, we have a woman president and vice president. Last year, Ana was president and I was vice president. This time around, I'm president and Sofia is vice president. We're showing a bigger appearance for women in the CS area, and wanting to do more things. But part of our club, part of our goal, is not just to help you tutor or program or something like that. We try and get all the CS people together outside of classes, and outside of school work, and socialize, get to know each other. It's been a great way to get to know people. Every year it seems like the officers seem to be getting better, and having the interests of the club members in mind. We got more involved with community service and did relay for life. Last year, we actually took charge of doing a trip for the computer science club. We were able to raise money throughout the year, and we went up to an advanced computing center.

At other institutions, students did not start formal extra-curricular clubs, but developed informal peer networks for support. Academically-inclined Latinas developed relationships with each other in the academic spaces of research groups, conferences, or classes. These networks provided personal support and encouragement. Leticia commented on an informal network of women in her computer science department.

> We know that within our group, we are there for the same reasons. We need to just help each other and find support when somebody may need tutoring or something you are struggling in. Or just like a conversation, a cup of coffee, in those types of things I really find support.

In the local figured worlds of their academic departments, some women began to organize their subjectivities around the issue of underrepresentation in computing, and thus enacted identities that valued being a Latina in computing. As Holland et al. (1998) theorized, the social interactions in the localized and temporal spaces of the computing clubs and informal peer networks gave voice to the lived experiences of Latinas in computing. In this process, some women gained a deeper understanding of the cultural, political, and historical landscape of Latinas in computing, and produced identities that contested the privileging of white and Asian male "hacker" culture in computer science. Much like the role models that they accessed through conferences who helped to foster their own shifts in identity, some of the Latina undergraduates served as "significant narrators" to younger students of local figured worlds in which they participated (Urrieta, 2007). In this way, the Latinas produced identities that recognized the intersectionality of race and gender within computing, and re-defined computer scientist as someone who values community, outreach, and mentoring over individualism and competition.

Conclusion

The Latinas in our study shared several common threads in their life history experiences that facilitated shifts in their identity within the professional and academic landscape of computing. The four life experiences related to professional empowerment were: 1) experiencing discrimination and oppression, 2) gaining a critical understanding of oppression through interactions with role models from similar backgrounds, 3) engaging in successful disciplinary performances, and 4) teaching or mentoring others. The Latinas in our study recounted many instances of bias and microaggressions. Yet participants' shift away from isolation and marginalization was fostered by interactions with Latina role models who situated the student's negative experiences within inequitable social, cultural, and historical legacies in computing. Participants began to see that their struggles were not isolated or unique. The role models served as catalysts to transform the student's understanding of their own experiences and helped to re-define their professional identities as Latinas in computing. Additionally,

successful disciplinary performances in research and at conferences were pivotal in conferring status to the women within their departments and the broader landscape of computing. These two experiences—interactions with role models and successful disciplinary performance—helped the women to see themselves, and to be seen, as competent computer scientists. Finally, participants enacted their professionally empowered identities by teaching and mentoring others. These experiences represent shifts away from an individualistic, isolated experience of computing to one defined by reflection, community, and teaching.

The Latinas in this study challenged the taken-for-granted notion of computing as a competitive, individualistic enterprise that is the domain of white or Asian, male hackers. Some of the students crafted identities as engaged, community-oriented computer scientists and enacted these identities in their departments and local communities. This study provides a glimpse into the way in which interactions and practices can foster, or impede, identification with the discipline for underrepresented students in computing. The research literature in STEM education has extensively explored the positive effects of role models but has rarely examined the ability of role models to illuminate social and cultural disparities and empower novices. However, through their interactions with role models and successful disciplinary performances, participants re-defined what it means to be a computer scientist and shaped computer science identities that value community, outreach, and cooperation over individualism and competition.

References

Barker, L.J. & Aspray, W. (2006). The state of research on pre-college experiences of girls with information technology. In *Women and information technology: Research on the reasons for under-representation*. Cohoon, J & Aspray, W. (eds.), 1-53. Cambridge, MA: MIT Press.

Beyer, S., Raines, K., & Haller, S. (2004). Deterrent to women taking computer science courses. *IEEE Society and Technology. 23*, 21-28.

Bonous-Hammarth, M. (2000). Pathways to success: Affirming opportunities for science, mathematics, and engineering majors. *The Journal of Negro Education*, *69*, 1/2, 92-111.

Carlone, H. & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, *44*(8), 1187-1218.

Clewell, B.C. & Braddock, J.H (2000). Influences on minority participation in mathematics, science, and engineering. In *Access denied: race, ethnicity, and the scientific enterprise*. Campbell, G., Denes, R. & Morrison, C. (eds). pp. 89-137. New York, NY: Oxford University Press.

Cole, D. & Espinoza, A. (2008). Examining the academic success of Latino students in science, technology, engineering and mathematics (STEM) majors. *Journal of College Student Development*. 49(4), 285-300.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, *43*, 1241-1299.

European Commission (2009). *She Figures 2009: Statistics and indicators on gender equality in science*. Retrieved on 1/22/13 from http://ec.europa.eu/research/science-society/document_library/pdf_06/she_figures_2009_en.pdf

Fontana, A. & Frey, J.H. (2000). The interview: From structured questions to negotiated text. In N.K. Denzin & Y.S. Lincoln, (Eds.), *Handbook of Qualitative Research* (2nd ed.), 645-672. Thousand Oaks: Sage.

Gasbarra, P., & Johnson, J. (2008). *Out Before the Game Begins: Hispanic Leaders Talk about What's Needed to Bring More Hispanic Youngsters into Science, Technology and Math Professions*. A Public Agenda Report Prepared for America's Competitiveness: Hispanic Participation in Technology Careers Summit. Palisades, NY.

Gloria, A., Castellanos, J., Lopez, A.G., & Rosales, R. (2005). An examination of academic nonpersistence decisions of Latino undergraduates. *Hispanic Journal of Behavioral Sciences*, *27*(2), 202-223.

Holland, D., Lachicotte, W., Jr., Skinner, D., & Cain, C. (1998). *Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.

Hurtado, S., Carter, D.F., & Spuler, A. (1996). Latino student transition to college: Assessing difficulties and factors in successful college adjustment. *Research in Higher Education*, 37(2), 135-157.

Margolis, J. (2008). *Stuck in the Shallow End: Education, Race, and Computing*. Cambridge, MA: MIT Press.

Margolis, J. & Fisher, A. (2002). *Unlocking the Clubhouse: Women in Computing*. Cambridge, MA: MIT Press.

National Center for Education Statistics (NCES), (2010-2011). *Digest of Education Statistics*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Center for Women and Information Technology (NCWIT) (2004). *Revolutionizing the face of technology*. Retrieved on 2/5/13 from www.ncwit.org.

National Science Board (NSB). 2012. *Science and Engineering Indicators Digest 2012.* Arlington VA: National Science Foundation (NSB 12-02).

Oseguera,L., Locks, A., & Vega, I. (2009). Increasing Latina/o Students' Baccalaureate Attainment: A focus on Retention. *Journal of Hispanic Higher Education*, *8*, 23 – 53.

Santiago, D, & Treindl, T. (2009) *Taking Stock: Higher Education and Latinos*. Excelencia in Education, Washington, DC: December, 2009.

Seymour, E. & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder: Westview Press.

Spradley, J. P. (1980). Participant Observation. New York: Holt, Rinehart, and Winston.

Stevens, R., O'Connor, K., Garrison, L., Jocuns, A., & Amos, D.M. (2008). Becoming an engineer: Toward a three dimensional view of engineering learning. *Journal of Engineering Education*, *97*(3), 355-368.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. *American Psychologist*, *62*(4), 271-286.

Tyson, W., Lee, R., Borman, K. & Hanson, M.A. (2007). Science, Technology, Engineering, and Mathematics (STEM) Pathways: High School Science and Mathematics Coursework and Postsecondary Degree Attainment. *Journal of Education for Students Placed at Risk (JESPAR), 12*(3), 243-270.

UNESCO Institute of Statistics (2009). *UIS Fact Sheet*. No 2, October, 2009. Retrieved on 1/22/13 from http://www.bpf.cas.cn/zlyj/200910/P020091023325511624098.pdf

United States Census (2011). *American Fact Finder*. Retrieved on 9/15/13 from http://www.factfinder2.census.gov.

Urrieta, L. Jr. (2007). Identity production in figured worlds: How some Mexican Americans become Chicana/o activist educators. *Urban Review*, *39*(2), 117-144.

Wajcman, Judy (2010) Feminist theories of technology. *Cambridge journal of economics*, *34* (1). 143-152.

Warschauer, M., Knobel, M. & Stone, L. (2004). Technology and Equity in Schooling: Deconstructing the Digital Divide. *Educational Policy*, *18*(4), 562-588.

Identifying Transfer of Inquiry Skills across Physical Science Simulations using Educational Data Mining

Michael Sao Pedro, Worcester Polytechnic Institute, 100 Institute Rd. Worcester MA 01609, mikesp@wpi.edu
Yang Jiang, Luc Paquette, Ryan S. Baker, Teachers College, 525 W. 120th St. New York, NY 10027
Email: yang.jiang@tc.columbia.edu, luc.paquette@gmail.com, ryan@educationaldatamining.org
Janice Gobert, Worcester Polytechnic Institute, 100 Institute Rd. Worcester MA 01609, jgobert@wpi.edu

Abstract: Students conducted inquiry using simulations within a rich learning environment for 4 science topics. By applying educational data mining to students' log data, assessment metrics were generated for two key inqury skills, testing stated hypotheses and designing controlled experiments. Three models were then developed to analyze the transfer of these inquiry skills between science topics. Model one, Classic Bayesian Knowledge Tracing, assumes that either complete transfer of skill occurs or no transfer occurs; model two (BKT-PST), an extension of BKT, assumes partial transfer and tests that assumption; and model three, a variant of BKT-PST, assumes no transfer and tests this assumption. An analysis of models one and two suggest that transfer of these inquiry skills across topics did occur. This work makes contributions to methodological approaches for measuring fine-grained skills using log files, as well as to the literature on the domain-specificity vs. domain-generality of inquiry skills.

Introduction

Science educators and reseachers agree that inquiry skills are critical to science literacy (NRC, 2011; Kuhn, 2005). To cultivate skills, some researchers have developed interactive, computer-based activities like simulations and microworlds (e.g. Quellmalz et al., 2009). A benefit of these activities is that they yield rich log data which can be leveraged for fine-grained performance assessment (Pellegrino et al., 2001; Mislevy et al., 2012). Though promising, assessment is still challenging because inquiry is multi-faceted, and manifests itself over time in complex ways (Williamson et al., 2006). Some are addressing these challenges using Educational Data Mining (EDM) to automatically assess specific skills (e.g. Sao Pedro et al., 2013a; Baker & Clarke-Midura, 2013; Ketelhut et al., 2013), Such techniques have potential to not only provide teachers and students real-time feedback about skill progress, but also to contribute to the field's understanding of inquiry learning.

In this paper, we use existing EDM models for evaluating data collection inquiry skills (Sao Pedro et al., 2012, 2013a) to build new models that identify skill transfer across several science topics. We focus on these skills because they support the development of other sense-making skills such as interpreting data and warranting claims (e.g. Kuhn, 2005), and because students have difficulty with these (de Jong & van Joolingen, 1998). Inquiry skills will be particularly valuable if they can transfer (Thorndike & Woodworth, 1901; Singley & Anderson, 1989), but it has been suggested that skills are tightly tied to the domain in which they are learned (van Joolingen et al., 2007), and thus may not transfer to new topics. However, other researchers have found evidence that inquiry skills can transfer and have a domain-general component (Glaser et al., 1991; Harrison & Schunn, 2004), or that content knowledge and inquiry skills co-develop (Kuhn et al., 1992; Kuhn & Pease, 2008). Though impressive, these studies had relatively small sample sizes and conflated data analysis skills with experimental design skills, skills unpacked in the present study. Our approach builds on our prior research (Sao Pedro et al., 2013c) in which we extended Bayesian Knowledge Tracing (Corbett & Anderson, 1995) to evaluate transfer of two data collection inquiry skills across two science topics. In particular, we address inquiry skill transfer at a larger scale with more students and across more science topics than seen in prior work.

Methodology

Participants

Participants were 299 eighth grade students from five middle schools in suburban Central Massachusetts who conducted inquiry across at least two science topics within Inq-ITS.

Materials: Inq-ITS Learning Environment Physical Science Inquiry Activities

Inq-ITS (**Inq**uiry **I**ntelligent **T**utoring **S**ystem, Gobert et al., 2012) is a web-based virtual science lab environment that automatically assesses students' inquiry skills (NRC, 2011). In this environment, students conduct inquiry with interactive simulations aligned to middle school Physical, Life, and Earth Science content described in the Massachusetts curricular frameworks, and inquiry support tools. In this paper, we focus on inquiry activities for four Physical Science topics: Phase Change (Figure 1), Free Fall Energy (Figure 2), Free Fall Speed, and Liquid Density. In a typical inquiry activity, students are first presented with a driving question.

For example, in a typical activity for Phase Change, students are asked to determine if one factor (e.g. size of container or amount of ice) affects specific outcomes (e.g. boiling point of water). Then, they conduct a semi-structured scientific inquiry process to address the goal: First, they articulate a hypothesis to be tested using a hypothesis widget with pulldown menus. Next, students collect data to try and test their hypothesis with a simulation. Students are required to run at least one trial before continuing. Once they finish running trials, they analyze their data by forming an argument (similar to hypothesizing) and selecting trials as evidence. A key aspect of this system is that activities provide performance assessment metrics on students' inquiry skills. Assessment of inquiry is based on the processes a student follows while experimenting, and the work products s/he creates using the support widgets.

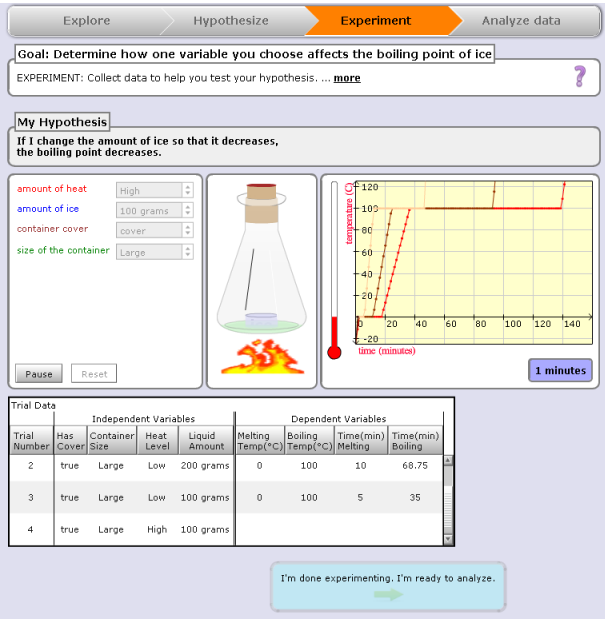


Figure 1. Phase Change simulation

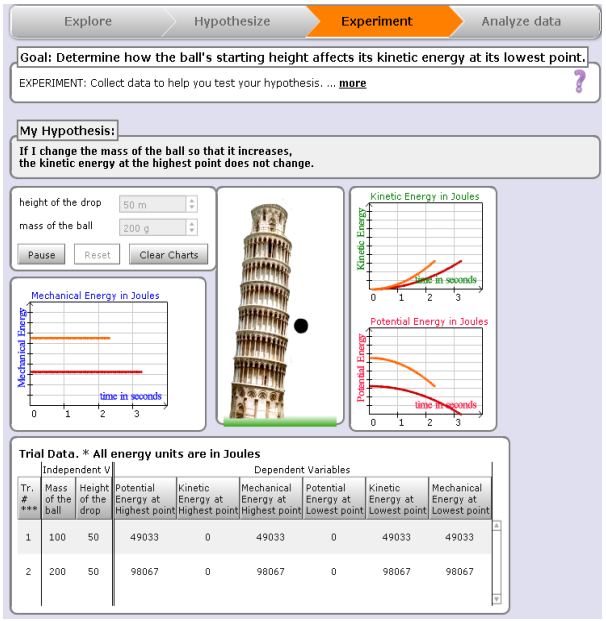


Figure 2. Free Fall Energy simulation

Procedure

Throughout the 2011-2012 school year, students at the five partner schools participated in inquiry within Inq-ITS. We coordinated with teachers regarding which activities would be used and when. Table 1 shows the specific activities chosen by each school and the order they were administered.

Each science topic included between 3 and 5 activities and were administered over two class periods of about 45 minutes each. Over the year, students completed between 2 and 4 sets of activities. The time delay between activity sets varied between schools, according to the respective teacher's pedagogical decisions (see Table 1). For example, at school 4, two science topics were done without any time gap (Free Fall Speed, then Free Fall Energy); at school 5, Free Fall Speed and Free Fall Energy were assigned 3.5 months apart. As students worked, Inq-ITS automatically logged all students' interactions, and automatically assessed their inquiry skills, as described in the next section. Unlike other Inq-ITS activities that provide personalized support (Sao Pedro et al., 2013c), students did not receive any explicit feedback on their inquiry processes or work products in the activities used in this study.

Evaluating Students' Data Collection Skills within Activities

Our work focuses on two data collection skills: designing controlled experiments, and collecting data to test hypotheses (Figures 1 and 2). Students design controlled experiments when they generate trials that make it possible to infer how changeable factors affect outcomes. This skill is related to the Control of Variables Strategy (CVS; cf., Chen & Klahr, 1999) that focuses on creating a single, contrastive and controlled experiment (a single pair of sequential trials). Unlike CVS, designing controlled experiments takes into consideration all a student's trials overall to determine whether a student demonstrates this skill (Sao Pedro et al. 2013a). The second skill, collecting data to test a hypotheses, is demonstrated when a student collects data that can support or refute an explicitly stated hypothesis. We track this in addition to designing controlled experiments because: 1) students may attempt to test their hypotheses with confounded designs, or may design controlled experiments for a hypothesis not explicitly stated; and 2), skill at testing hypotheses may be indicative of a student's successful planning and monitoring of their inquiry (de Jong, 2006).

Our process skills assessment is based on students' actions taken while collecting data with the simulation, and we evaluate whether students design controlled experiments and collect data to test their hypothesis using a combination of data-mined detectors and knowledge-engineered rules (Sao Pedro et al., 2013a,b). Our data mining approach accounts for "corner" cases when students do not conduct their inquiry in lock-step fashion, unlike other approaches that require sequential trials as demonstration of CVS (e.g. McElhaney & Linn, 2010). The goodness and generalizability of data mined detectors also can be determined by testing how well they can predict skill for students who were not used to build the detectors, e.g., we conducted extensive validation tests to show that these detectors agree with expert judgments of inquiry skill performance across our physical science activities (Sao Pedro et al., 2013b,c; Gobert et al., 2013), and new student populations (Sao Pedro et al., 2013c). The detectors are the backbone for generating models of skill transfer across topics, discussed in the next section.

Table 1: Topic order for each school, time delay between activities, and number of participants who conducted inquiry in each pair of topics.

Simulation Topic Pair	School	Delay Between Topics	Number of Participants	Error Rate (% students not demonstrating skill)		
				Last Attempt 1st topic	First Attempt 2nd Topic	Last Attempt 2nd Topic
PhCh → FF Energy	1,2,3	2-3 weeks	140	CtrlExp: 21.4% TestHyp: 21.4%	CtrlExp: 19.3% TestHyp: 14.3%	CtrlExp: 14.3% TestHyp: 15.7%
Density → FF Speed	4	5 weeks	33	CtrlExp: 69.7% TestHyp: 18.2%	CtrlExp: 24.2% TestHyp: 24.2%	CtrlExp: 12.1% TestHyp: 12.1%
FF Speed → FF Energy		no delay	31	CtrlExp: 12.9% TestHyp: 12.9%	CtrlExp: 16.1% TestHyp: 12.9%	CtrlExp: 6.5% TestHyp: 6.5%
FF Energy → PhCh		3 weeks	31	CtrlExp: 12.9% TestHyp: 12.9%	CtrlExp: 9.7% TestHyp: 3.2%	CtrlExp: 9.7% TestHyp: 9.7%
FF Energy → FF Speed	5	14 weeks	64	CtrlExp: 62.5% TestHyp: 57.8%	CtrlExp: 40.6% TestHyp: 39.1%	CtrlExp: 42.2% TestHyp: 42.2%

Developing Models of Transfer to Track Students' Performance across Topics

We model student knowledge and estimate the probability that students are transferring science inquiry skill between topics using Bayesian Knowledge Tracing (BKT, Corbett & Anderson, 1995). BKT is a two-state model (in technical terms, a Hidden Markov Model or simple Dynamic Bayesian Network) that estimates whether student knows a specific latent skill, based on the student's past history of observed performance on that skill. Here, we use BKT to estimate if students know how to design controlled experiments and how to collect data to test a hypothesis (cf. Sao Pedro et al., 2013a,c). The observable performance is whether a student actually demonstrated skill, determined by the detectors discussed previously. BKT has been widely and successfully used to model student knowledge in various intelligent tutoring systems, including the widely-used Cognitive Tutor (Corbett & Anderson, 1995) and ASSISTments systems (Pardos & Heffernan, 2010). BKT performs equivalently to or better than competing approaches (Gowda et al., 2011), and has been extended to support analysis of the nature of student learning (e.g. Beck et al., 2008; Sao Pedro et al., 2013c).

In the classic BKT framework, it is assumed that a skill is either known or not known, and that there is a certain probability of each. Students demonstrate an inquiry skill when (1) they already know the skill and they do not make a slip (a careless mistake); or when (2) they do not know the skill but guess how to do it correctly. The model is defined by a set of four parameters: $P(L_0)$, the probability that the skill is already known before the first opportunity to use it; *T,* the probability that the skill will be learned at each opportunity to use it (classical BKT does not include forgetting, though many extensions do); *G,* the probability that a student will guess and demonstrate the skill despite not knowing it; and *S*, the probability that the student will slip and make a mistake despite knowing the skill. In classical BKT, the four parameters are assumed to be the same for all students (many variants on BKT relax this constraint as well).

Using these parameters, the classic BKT model can incrementally calculate the likelihood $P(L_n)$ that a student knows a skill, such as how to design controlled experiments, after the student finishes their n^{th} attempt practicing the skill ($Prac_n$) in an inquiry activity. It can also estimate the likelihood that a student will demonstrate a skill *before* they begin their inquiry in the *n*th attempt, $P(Prac_n = True)$ using the prior estimate of knowledge, $P(L_{n-1})$. The equations for computing these two estimates are as follows:

$$PLn = PLn-1 Pracn + 1 - PLn-1 Pracn * T, \text{ where}$$

$$PLn-1 Pracn = False = PLn-1 * 1 - SPLn-1 * 1 - S + 1 - PLn-1 * G$$

$$PLn-1 Pracn = True = PLn-1 * SPLn-1 * S + 1 - PLn-1 * 1 - G$$

$$PPracn = True = PLn-1 * 1 - S + 1 - PLn-1 * G$$

One assumption of the Classic BKT model that is relevant to the present work is that it assumes either that complete transfer of skill occurs or no transfer occurs (cf. Sao Pedro et al., 2013c). That is, in Classic BKT, full transfer can be assumed by treating two skills as the same skill (e.g., designing controlled experiments is the same skill whether it is in Phase Change or Density); no transfer is assumed by treating the skill as a separate, independent skill within each topic (e.g. designing controlled experiments in Phase Change and designing controlled experiments in Density are different skills). Since we believe that the acquisition of inquiry skills is richer than this, rather than make either assumption, we developed an extension to BKT that aims to capture the possibility of partial transfer of skill (cf. Singley & Anderson, 1989) across science topics, the BKT-PST model. Capturing partial transfer enables us to determine empirically whether transfer occurred and the degree to which it occurred across pairs of science topics.

BKT-PST: Accounting for Partial Transfer of Skills

The proposed BKT-PST model builds upon our prior work (Sao Pedro et al., 2013c) in which we extended BKT to account for partial transfer. In this work, we added two components in the model to adjust the likelihood of knowing a data collection skill, $P(L_n)$, in a new science topic. The first was an observable $Topic_Switch_n$ = {True, False} to indicate when the student begun a new set of inquiry activities for a different science topic. The second was a degradation parameter, $k \in (0.0, 1.0)$ that lowers the likelihood of knowing the skill by a constant factor k when switching science topics. The k parameter captures that students may not readily know to apply (transfer) the same data collection skills within different simulations (cf. Singley & Anderson, 1989). We believe, though, that the original approach may not accurately model transfer. Though $k = 1$ in this model accurately models full transfer (the estimate $P(L_n)$ does not get degraded when the topic switches), $k = 0$ would predict with certainty that the student would have no skill at all, degrading $P(L_n)$ to be 0. Thus, for low values of k the model may be too strict.

The BKT-PST model has the same Bayesian Network topology as our prior work (Sao Pedro et al., 2013c), but instead we change how the k parameter impacts the estimate of $P(L_n)$. In BKT-PST, the k parameter represents the percentage of learning accumulated within the first science topic that is transferred to the second topic. So, when $Topic_Switch_n$ = True, the likelihood that students know the skill before the second science topic $P(L_n)$ is equal to the sum of the initial latent skill, $P(L_0)$, and the learning that is transferred, k*($P(L_{n-1} \mid Prac_n)$ - $P(L_0)$). The modified equations to compute $P(L_n)$ for BKT-PST become:

$$PLnTopic_Switch_n=True=PST+1-PST*T, \text{ with } PST= PL0+k*PLn-1|Pracn-PL0$$

$$PLnTopic_Switch_n=False=PLn-1Pracn+1-PLn-1Pracn*T$$

If full transfer is assumed ($k = 1$), BKT-PST behaves the same way as the classic BKT model and indicates that a student's latent skill does not degrade for a new topic. When $k = 0$, $P(L_n)$ returns back to the original estimate of initial knowledge, adjusted for the possibility of learning from the practice attempt, a more realistic assumption. In other words, mathematically when $k = 0$, $P(L_n \mid Topic_Switch_n = True) = P(L_0) + (1 - P(L_0))*T$.

Though BKT-PST may better represent transfer, it is worth noting that it has an important limitation for a somewhat uncommon special case. Take a student who fails to demonstrate the skill completely on all attempts (*n-1* attempts) in the first science topic. After observing all these failures, the likelihood of knowing the skill, $P(L_{n-1})$, will be less than $P(L_0)$. In this case, for sufficiently low values of *k,* the PST computation will be *larger* than $P(L_{n-1} \mid Prac_n)$, the updated estimate of $P(L_{n-1})$ after observing the performance $Prac_n$. In other words, the BKT modification could yield an increase when switching topics, the opposite of our assumption of degradation after switching topics. For unusual cases like this, BKT-PST may not be an ideal model. In our study, this special case occurred on 18.06% and 26.42% of topic switches for the skill of designing controlled and testing stated hypotheses, respectively.

Model Fitting

We employed a brute force grid search approach (Baker et al., 2010), a standard approach for fitting BKT models, to determine the value of each set of parameters for our three models. In order to find the best-fitting parameters, all potential parameter combinations in the search space were tried at a grain-size of 0.01 for each skill per simulation pair. The best set of parameters is the one that yields the lowest sum of squares residual (SSR) between the likelihood of demonstrating skill, $P(Prac_n = True)$, and the actual data. The values of Guess parameter G and Slip parameter S were bounded to be below 0.5 to avoid "model degeneracy" (Baker et al., 2008), where a model may estimate that the student has a lower probability of knowing $P(L_n)$ after observing the student demonstrate the skill. All other variables were allowed to have values from 0.01 to 0.99. For the previously found best parameter set, the same brute force search process was repeated around these parameters at a grain-size of 0.001 to find a tighter fit. For the "no transfer" BKT-PST model, we applied the classic BKT model on the data from activities in both science topics to calculate overall $P(L_0)$, before using brute force grid search strategy again to calculate the other three parameters. This was done in order to avoid the model from accounting for transfer of skills across science topics by increasing the initial learning probability $P(L_0)$.

Results

As previously mentioned, we applied three models that assume, respectively, full transfer, no transfer, or partial transfer to evaluate students' mastery and transfer of science inquiry skill between pairs of science topics (e.g. between Phase Change and Free Fall Energy, Density and Free Fall Speed, etc.). This is done in two ways. First, we fit and compare the parameters of three models: Classic BKT (Corbett & Anderson, 1995) that assumes full transfer, our new BKT-PST model that empirically estimates partial transfer (BKT-PST with $k >= 0$), and a model that assumes data collection skills do not transfer across science topics (BKT-PST with $k = 0$). Then, we conduct a more stringent test, comparing whether Classic BKT and BKT-PST predicts student performance better than the no transfer model. If the Classic BKT or BKT-PST models fit student performance data better than the no transfer model, it implies that transfer occurred.

To get a sense of student performance across the activities, we first conducted a descriptive analysis by computing error rates (% students who fail to demonstrate the skills during a practice opportunity) at three key points at each topic pair where transfer (or the lack thereof) can be seen: 1) at the last practice opportunity of the first science topic, 2) at the first practice opportunity for the second science topic, and 3) at the last practice opportunity of the second topic. In order for the calculated error rates to be indicative of transfer, we expect error rates to be constant or decreasing at each point. As shown in Table 1, almost all error rates fit this criteria. There were only two exceptions where error rates slightly increased – from 18.2% to 24.2% when transferring the skill of testing hypotheses from "Density" to "Free Fall Speed" activities, and from 12.9% to 16.1% when transferring the skill of designing controlled experiments from "Free Fall Speed" to "Free Fall Energy" activities. In addition, the error rates were always smaller for the last practice opportunity of the second topic than the last practice opportunity of the first topic. Though we report on these three specific points, we note that *all* of the students' activities were used to construct and test the models.

Interpreting Parameters of the Predictive Models

Partial transfer of science inquiry skills across topics was captured by the linear transfer factor k for "transfer" models built for five science topic pairs (see Table 2). The high value of k for almost all skills and topic pairs (between 0.839 and 0.990) suggests close to full skill transfer across science topics. The only exception was for the skill of testing hypotheses for students in the "Free Fall Energy" and "Free Fall Speed" topic pair for which transfer was poor. In this special case, both the learning rate ($T = 0.001$), and transfer of what they learned in the first science topic to the second topic ($k = 0.341$) were low. But overall, these finding suggests these inquiry skills transfer between science topics in Inq-ITS, replicating earlier findings (Sao Pedro et al., 2013c).

We can understand these models better by looking at the four remaining parameters of the BKT-PST models ($P(L_0)$, *S, G, T*), and comparing these parameters to those in the classic BKT models. Because the BKT-PST model behaves exactly the same as the classic BKT model when the linear transfer factor k is equal to 1, and the k values were high across the different science topic pairs, we would expect the four remaining parameters to be very similar between models. The highly similar parameters obtained when comparing classic BKT models and the "transfer" models (Table 2) meet our expectations and indicate that transfer occurred. As such, we can conclude that students were able to apply what they had learned about data collection skill from one science topic to another with very little degradation of the skill. For the Free Fall Energy to Phase Change pair, we noticed that the Guess parameter (G) for both the classic BKT model and the "skill degradation model" hit its 0.5 boundary for designing controlled experiments skill, indicating that students who did not know the skill were as likely to get the question correct by guessing than they were to get it incorrect. More research will be needed to determine why this occurred.

Comparing Models' Overall Predictive Capability

To test and compare the how well the three BKT models performed in accurately tracking the development of each inquiry skill, we conducted six-fold student-level cross validation for all science topic pairs to determine which models hold better predictive performance in predicting skill demonstration. Specifically, we stratified students randomly into six folds per skill per science topic pair and trained and tested the models' performance by comparing the estimated $P(Prac_n = True)$ with the actual student performance at time n. This cross-validation process ensures that the models can be generalized to other groups of students beyond those whose data were used to train the models originally. Model goodness was determined by computing A', which is the probability that the model will be able to distinguish practice opportunities in which the skill is and is not demonstrated. A' was used because it is an appropriate metric to use when using predictions with a confidence value to predict a binary variable (Fogarty et al., 2005). An A' value of 0.5 implies chance-level performance, and one of 1.0 indicates perfect performance.

Overall, the "transfer" models and the classic BKT models showed similar model goodness with A' values ranging from 0.512 to 0.870 for designing controlled experiments skill, and values ranging from 0.575 to 0.900 for testing stated hypotheses skill (see Table 2). All the A's per skill per simulation pair are above the 0.5 chance level. The "transfer" model for testing stated hypotheses in the "Free Fall Energy" to "Free Fall Speed"

pair performed slightly better than the classic BKT model and an unusually low A' (0.512) was observed for the classic BKT model of designing controlled experiments in the "Free Fall speed" to "Free Fall Energy" pair. Further investigation showed that one of the training folds for "Free Fall speed" to "Free Fall Energy" pair yielded a very high learning rate (T = 0.990), which causes the model to immediately update its P(L_n) estimate to be near 1.0. The low A' is consistent with this type of degenerate model. It is also worth noting that the A' obtained for the "no transfer" BKT models are lower than those of the corresponding BKT-PST models. The fact that the BKT-PST models performed better than the "no transfer" models at predicting the students' skills, combined with the high values obtained for the linear transfer factors k supports our hypothesis that science inquiry skill transfers between two science topics.

Table 2: Parameter values for four BKT models across all science topic pairs.

Topic Pair	Model	Skill	$P(L_0)$	G	S	T	K	A'
Phase Change → Free Fall Energy (n = 140) Schools 1,2,3	Classic BKT	CtrlExp	0.621	0.138	0.053	0.142		0.870
		TestHyp	0.645	0.145	0.036	0.130		0.895
	No Transfer	CtrlExp	0.621	0.168	0.035	0.173		0.836
		TestHyp	0.645	0.177	0.018	0.150		0.860
	Transfer	CtrlExp	0.617	0.142	0.052	0.144	0.990	0.867
		TestHyp	0.615	0.149	0.017	0.148	0.904	0.891
Free Fall Energy → Phase Change (n = 31) School 4	Classic BKT	CtrlExp	0.839	0.131	0.005	0.259		0.829
		TestHyp	0.879	0.161	0.002	0.119		0.897
	No Transfer	CtrlExp	0.839	0.336	0.001	0.259		0.800
		TestHyp	0.879	0.171	0.003	0.169		0.834
	Transfer	CtrlExp	0.839	0.131	0.007	0.259	0.990	0.832
		TestHyp	0.879	0.171	0.002	0.119	0.990	0.900
Density → Free Fall Speed (n = 33) School 4	Classic BKT	CtrlExp	0.147	0.158	0.174	0.323		0.731
		TestHyp	0.489	0.001	0.06	0.356		0.831
	No Transfer	CtrlExp	0.147	0.383	0.001	0.092		0.606
		TestHyp	0.489	0.181	0.003	0.372		0.739
	Transfer	CtrlExp	0.148	0.159	0.173	0.323	0.990	0.730
		TestHyp	0.489	0.001	0.060	0.357	0.990	0.829
Free Fall Speed → Free Fall Energy (n = 31) School 4	Classic BKT	CtrlExp	0.531	0.500	0.001	0.279		0.512
		TestHyp	0.661	0.371	0.001	0.229		0.757
	No Transfer	CtrlExp	0.531	0.500	0.001	0.368		0.611
		TestHyp	0.661	0.471	0.001	0.139		0.599
	Transfer	CtrlExp	0.531	0.500	0.001	0.279	0.969	0.672
		TestHyp	0.692	0.321	0.001	0.239	0.839	0.760
Free Fall Energy → Free Fall Speed (n = 64) School 5	Classic BKT	CtrlExp	0.505	0.001	0.354	0.297		0.642
		TestHyp	0.287	0.33	0.174	0.087		0.575
	No Transfer	CtrlExp	0.505	0.176	0.253	0.079		0.632
		TestHyp	0.287	0.378	0.072	0.001		0.573
	Transfer	CtrlExp	0.506	0.001	0.353	0.296	0.990	0.633
		TestHyp	0.429	0.318	0.164	0.001	0.341	0.593

Discussion and Conclusions

In this paper, we leveraged Educational Data-Mined models to investigate whether two data collection inquiry skills, designing controlled experiments and testing stated hypotheses, transferred across four physical science simulations in Inq-ITS (Gobert et al., 2012). To empirically test for transfer, we developed two different models based on Bayesian Knowledge Tracing (Corbett & Anderson, 1995). Each makes different presuppositions about the likelihood of transfer occurring. The first, Classic BKT (Corbett & Anderson, 1995), assumes either *complete transfer,* or *complete skill independence*. The second model, BKT-PST, captures partial transfer of skill. BKT-PST assumes that inquiry skills are more nuanced in their acquisition and transfer, and that they are likely to be honed more gradually. We determined whether transfer occurred between topics by comparing the BKT-PST transfer model to a BKT-PST model with the assumption of no transfer, which posits that skills are tied to the domain in which they are learned (cf. van Joolingen et al., 2007). Our results indicated that both skills

transferred across nearly all the pairs of the physical science topics tested. This was demonstrated by the BKT-PST transfer parameter having very high values for both inquiry skills. In addition, we found that the BKT-PST model better captured student performance than the BKT-PST model with no transfer assumed in 5 of the 6 topic pairs for both inquiry skills, increasing our confidence that transfer occurred.

This paper makes two main contributions towards understanding of inquiry learning and scalable, performance-based assessment of inquiry. Our findings contribute to the understanding of domain-specificity vs. domain-generality of inquiry skills (Kuhn et al., 1992; Klahr & Nigam, 2004; van Joolingen et al., 2007) since they suggest that skills have some domain-general aspects. For example, once one knows how to design experiments, they can do so in a new domain to better understand a phenomena under investigation (e.g. Gobert et al., 2012). However, we note that all of the physical science simulations studied here have a similar, linear causal structure, which may have facilitated transfer. Skills may manifest themselves differently for simulations with more complex causal systems (Jacobson & Wilensky, 2006). In addition, we note that transfer was determined for activities solely within the learning environment. In the future, it will be beneficial to determine if the models can also predict skill knowledge on other tests external to the system in order to better understand how general these inquiry skills are (Klahr & Nigam, 2004; Baker et al., 2011).

This work also contributes to the literature on scalable, performance-based formative assessment of inquiry skills across domains. Our models explicitly capture skill transfer, and can be used to estimate students' performance and drive scaffolding in real-time (Sao Pedro et al., 2013c). We note that it is important to use metrics for inquiry skills that do not require that students conduct sequential experimental trials in lock-step fashion (e.g. McElhaney & Linn, 2010). Skill at designing controlled experiments can manifest itself multiple ways (Sao Pedro et al., 2013a), and a distinction needs to be made between students designing controlled experiments in unusual ways and students engaging in haphazard inquiry (cf., Buckley et al., 2010). One limitation of our BKT-PST model is that it cannot cleanly identify *what* causes transfer. There were substantial gaps in time between assessments during which teachers may have provided supports that helped students to acquire and transfer skills.

In closing, we note that developing science inquiry skills is a *necessary but not sufficient condition* for deep science learning. We believe that it is the application of these skills to science phenomena in rich meaningful ways that has the potential to result in deep conceptual learning. As such, being able to identify and track how these skills develop and transfer is crucial towards promoting rich skill development.

References

Baker, R., Clarke-Midura, J. (2013) Predicting Successful Inquiry Learning in a Virtual Performance Assessment for Science. In *Proc. of the 21st Int'l Conf. on User Modeling, Adaptation, and Personalization* (pp. 203-214).

Baker, R., Corbett, A., Aleven, V. (2008). More accurate student model through contextual estimation of slip and guess probabilities in Bayesian Knowledge Tracing. *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 406-415).

Baker, R., Corbett, A., Gowda, S., Wagner, A., MacLaren, B., Kauffman, L., et al. (2010). Contextual Slip and Prediction of Student Performance After Use of an Intelligent Tutor. In *Proceedings of the 18th Annual Conference on User Modeling, Adaptation, and Personalization* (pp. 52-63).

Beck, J., Chang, K., Mostow, J., & Corbett, A. (2008). Does Help Help? Introducing the Bayesian Evaluation and Assessment Methodology. In *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 383-394).

Baker, R., Gowda, S., & Corbett, A. (2011) Automatically Detecting a Student's Preparation for Future Learning: Help Use is Key. In *Proc. of the 4th Int'l Conf. on Educational Data Mining* (pp. 179-188).

Buckley, B.C., Gobert, J., Horwitz, P. & O'Dwyer, L. (2010). Looking inside the black box: Assessments and decision-making in BioLogica. *International Journal of Learning Technologies*, 5(2), 166 - 190.

Chen, Z., & Klahr, D. (1999). All Other Things Being Equal: Acquisition and Transfer of the Control of Variables Strategy. *Child Development,* 70(5), 1098-1120.

Corbett, A. T. & Anderson, J. R. (1995). Knowledge-Tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4,* 253-278.

de Jong, T. (2006). Computer Simulations - Technological advances in inquiry learning. *Science*, 312(5773), 532-533.

de Jong, T., & van Joolingen, W. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. Review of Educational Research, 68, 179-201.

Fogarty, J., Baker, R., & Hudson, S. (2005). Case Studies in the Use of ROC Curve Analysis for Sensor-Based Estimates in Human Computer Interaction. *Proc. of Graphics Interface* (pp. 129-136).

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1991). Scientific Reasoning Across Different Domains. In *Computer-based Learning Environments and Problem-Solving* (pp. 345-371).

Gobert, J., Sao Pedro, M., Baker, R., Toto, E., & Montalvo, O. (2012). Leveraging educational data mining for real time performance assessment of scientific inquiry skills within microworlds. *Journal of Educational Data Mining, 4*(1), 111-143.

Gobert, J., Sao Pedro, M., Raziuddin, J., & Baker, R. (2013). From Log Files to Assessment Metrics for Science Inquiry using Educational Data Mining. *Journal of the Learning Sciences, 22*(4), 521-563.

Gowda, S., Baker, R., Pardos, Z., Heffernan, N. (2011). The sum is greater than the parts: Ensembling student knowledge models in ASSISTments. In *Proc. of the KDD Workshop on KDD in Educational Data*.

Harrison, A., & Schunn, C. (2004). The Transfer of Logically General Scientific Reasoning Skills. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, (pp. 541-546).

Jacobson, M. J., & Wilensky, U. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *Journal of the Learning Sciences*, 15 (1), 11-34.

Ketelhut, D., Nelson, B., Sil, A., & Yates, A. (2013). Discovering what students know through data mining their problem-solving actions within the immersive virtual environment, SAVE Science. Presented at the American Educational Research Association, April 27-May 1, 2013, San Francisco, CA.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychological Science*, 15(10), 661-667.

Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.

Kuhn, D., & Pease, M. (2008). What Needs to Develop in the Development of Inquiry Skills? *Cognition and Instruction*, 26(4), 512-559.

Kuhn, D., Schauble, L., & Garcia-Mila, M. (1992). Cross-Domain Development of Scientific Reasoning. *Cognition and Instruction*, 9, 285-327.

McElhaney, K., & Linn, M. (2010). Helping Students Make Controlled Experiments More Informative. In *Proc. of the 9th International Conference of the Learning Sciences - Volume 1, Full Papers* (pp. 786-793).

Mislevy, R.J., Behrens, J.T., DiCerbo, K., & Levy, R. (2012). Design and discovery in educational assessment: Evidence centered design, psychometrics, and data mining. *J. of Educational Data Mining, 4*, 11-48.

National Research Council. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: The National Academies Press.

Pardos, Z. & Heffernan, N. (2010). Modeling individualization in a Bayesian Networks implementation of knowledge tracing. In *Proc. of the 18th Int'l Conf. on User Modeling, Adaptation and Personalization* (pp. 255-266).

Pellegrino, J., Chudowsky, N., and Glaser, R. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

Quellmalz, E., Timms, M., & Schneider, S., (2009). *Assessment of Student Learning in Science Simulations and Games*. Washington, DC: National Research Council Report, Washington, DC.

Sao Pedro, M., Baker, R., & Gobert, J. (2012). Improving Construct Validity Yields Better Models of Systematic Inquiry, Even with Less Information. In *Proceedings of the 20th Conference on User Modeling, Adaptation, and Personalization* (pp. 249-260).

Sao Pedro, M., Baker, R., Gobert, J., Montalvo, O., & Nakama, A. (2013a). Leveraging Machine-Learned Detectors of Systematic Inquiry Behavior to Estimate and Predict Transfer of Inquiry Skill. *User Modeling and User-Adapted Interaction, 23*, 1-39.

Sao Pedro, M., Baker, R., & Gobert, J. (2013b). What Different Kinds of Stratification Can Reveal about the Generalizability of Data-Mined Skill Assessment Models. *Proceedings of the 3rd Conference on Learning Analytics and Knowledge*, (pp. 190-194).

Sao Pedro, M., Baker, R., & Gobert, J. (2013c). Incorporating Scaffolding and Tutor Context into Bayesian Knowledge Tracing to Predict Inquiry Skill Acquisition. In *Proceedings of the 6th International Conference on Educational Data Mining*, (pp. 185-192).

Singley, M. & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard Univ. Press.

Thorndike, E.L. & Woodworth, R.S. (1901). The influence of improvement in one mental function upon the efficacy of other functions. *Psychological Reviews,* 8, 247-261.

van Joolingen, W., de Jong, T., & Dimitrakopoulout, A. (2007). Issues in Computer Supported Inquiry Learning in Science. *Journal of Computer Assisted Learning*, 23(2), 111-119.

Williamson, D., Mislevy, R., & Bejar, I. (2006). *Automated 0053coring of Complex Tasks in Computer-Based Testing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Acknowledgements

This research is funded by the National Science Foundation (NSF-DRL#0733286, NSF-DRL#1008649, and NSF-DGE#0742503) and the U.S. Department of Education (R305A090170 and R305A120778). Any opinions expressed are those of the authors and do not necessarily reflect those of the funding agencies.

Reactivation of Multimodal Representations and Perceptual Simulations for Meaningful Learning: A Comparison of Direct Embodiment, Surrogate Embodiment, and Imagined Embodiment

Saadia A. Khan, Teachers College, Columbia University
525 West 120th Street, Box 118, New York, NY 10027
Email: khan2@tc.columbia.edu

Abstract: Embodiment has been found to enhance learning and motivation. It is proposed that during embodied experiences, learners reactivate multimodal representations of previously stored memories related to objects and events, and the mental perceptual simulations learners construct during embodiment enhance their learning and motivation. This paper presents the findings of a study that investigated the effects of different types of embodiment on the learning and motivation of adult learners. The study compared four groups, Direct Embodiment, Surrogate Embodiment, Imagined Embodiment, and No Embodiment (control). The findings suggest that learners learn better and have higher motivation when they engage in embodied learning than when they experience no embodiment, and that role-playing as avatars in virtual environments and role-playing physically both enhance learning and motivation more than imagining actions and reading. The findings suggest that role-play can make learning more meaningful.

Introduction

According to the embodiment premise, cognition depends not just on the mind but also on the body and people's experiences of their bodies in action (Gibbs, 2007; Robbins & Aydede, 2009). Embodiment or using bodily movements to enact knowledge and concepts can help people learn (Lindgren & Glenberg-Johnson, 2013). Theories of embodied cognition propose that embodiment improves memory and comprehension since it involves the construction of mental perceptual simulations and the reactivation of multimodal representations initially stored in memory during a learner's previous experience with an object or event (Barsalou, 2008a, 2008b). Furthermore, it is proposed that embodiment can lead to positive transfer of learning since learners learn to imagine during an embodied learning experience, and they are then able to use their imagination in new learning situations (Black, Segal, Vitale, & Fadjo, 2012). Empirical research on imagination and embodiment is further supported by neuroimaging data that indicate that the same brain regions are activated when we perform an action and imagine an action (Buccino, Binkofski, et al., 2001; Buccino, Riggio, et al., 2005; Hauk & Pulvermuller, 2004; Pulvermuller, 2008). Research findings suggest that embodiment can enhance memory, comprehension, transfer, and motivation (Barab, Dodge, Thomas, Jackson, & Tuzun, 2007; Bianchi-Berthouze, Kim, & Patel, 2007; Black, Khan, & Huang, 2014; Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004; Khan, 2012, 2013; Khan & Black, 2013, in press; Metcalf, Dede, Grotzer, & Kamarainen, 2009; Metcalf, Kamarainen, Grotzer, & Dede, 2011; Noice & Noice, 2001; Scott, Harris, & Rothe, 2001).

How can embodiment be used effectively to make learning a more meaningful experience for learners? I propose that embodied instruction and learning should specifically involve role-play activities in which learners can construct imaginary worlds (Black, 2007) through bodily movements. Role-play, whether it is through physical movements or virtual movements or imagined movements, can be expected to provide learners with the opportunity to engage with the learning material at a deeper level. Role-playing involves (a) an internal process in which the role-player uses certain conceptual constructs (i.e., the character she is playing, the game world, and the story), and (b) sharing this internal process with others through external expression (Lankoski & Järvelä, 2012). Embodiment via role-play provides immersive learning experiences in which the learner can put herself in another person's shoes (whether this is a fictional character or a character from history) to better understand learning material. The learner gets the opportunity to embody the character and enact a story through movement, gestures, dialogue, and the expression of emotions. This enables learners to relate more to the learning material and it can make learning more enjoyable by introducing an element of play into mundane learning material. Role-play also allows learners and instructors to collaborate during embodied learning experiences, in which instructors and peers can provide scaffolding to learners.

Theoretical Background

In spite of a growing body of evidence suggesting that embodiment enhances learning and motivation, embodied cognition is still developing towards a unified theory (Shapiro, 2011). Embodied cognition theories that explain cognition via mental simulations and reactivation of multimodal representations include *Perceptual Symbol Systems*, the *Indexical Hypothesis* and *Basic Systems Theory*. According to *Perceptual Symbol Systems* (Barsalou, 2008a, 2008b), whenever we experience some object or event, we store the memory of that object or

event as multimodal representations. When we need to recall something about the object or event at a later time, we reactivate these multimodal representations as mental simulations of the object or event. *Basic Systems* also proposes that we simulate the multimodal components (including vision, audition, action, space, affect, and language) of a complex memory during retrieval (Barsalou, 2008a). The *Indexical Hypothesis* (Glenberg, 2008; Glenberg et al., 2004), which is related to *Perceptual Symbol Systems*, proposes that language comprehension involves (a) indexing words and phrases to objects in the environment or to perceptual symbols, (b) deriving affordances (Gibson, 1979) from the objects, and (c) combining the affordances according to syntax to produce a coherent simulation (Glenberg et al., 2004). The *Indexical Hypothesis* also corresponds with the *Ideomotor Theory* developed by Prinz about ideomotor mapping or forming learned associations between actions and their effects or between desired effects and actions (Glenberg, 2008). Boroditsky and Prinz (2008) have combined *Perceptual Symbol Systems* and *Indexical Hypothesis* to propose that two *Input Streams* are involved in cognition. According to them, people receive information from both perception and language and information is combined from these two input streams.

Black et al. (2012) have provided a theoretical framework that contributes to embodied cognition theory by specifically focusing on how embodiment can be used to deliver instruction. Their *Instructional Embodiment Framework* defines different types of embodiment that can be used for instruction and learning (see Figure 1). The framework divides embodied instruction into two main types of embodiment: Physical Embodiment and Imagined Embodiment. Physical Embodiment is further divided into Direct Embodiment (i.e., physically moving to perform actions during learning), Surrogate Embodiment (i.e., using a deputy to perform actions, such as controlling the actions of a virtual agent or avatar), and Augmented Embodiment (e.g., using touch devices and augmented reality). Imagined Embodiment is divided into Explicit (i.e., learners are explicitly instructed to imagine while learning) and Implicit (i.e., learners are not explicitly instructed to imagine while learning and imagining takes place at an implicit, latent level). According to Black (2007), imagination plays an important role in enhancing understanding and comprehension and constructing imaginary worlds during a learning activity can improve learners' memory and comprehension.

The *Instructional Embodiment Framework* proposes that embodied instruction is more effective as an instructional strategy than traditional non-embodied instruction since embodiment enables learners to construct mental perceptual simulations. Since mental perceptual simulations involve more than one modality, they can enable us to ascribe a deeper meaning to our experience making learning more meaningful. Based on this framework, a number of research studies suggest that learners learn better when instruction includes embodied learning experiences, and that embodied learning is further enhanced when there is a higher level of embodiment than when there is a lower level of embodiment or no embodiment (Black et al., 2012; Khan, 2012; Khan & Black, 2013, in press; Lu, Kang, Huang, & Black, 2011).

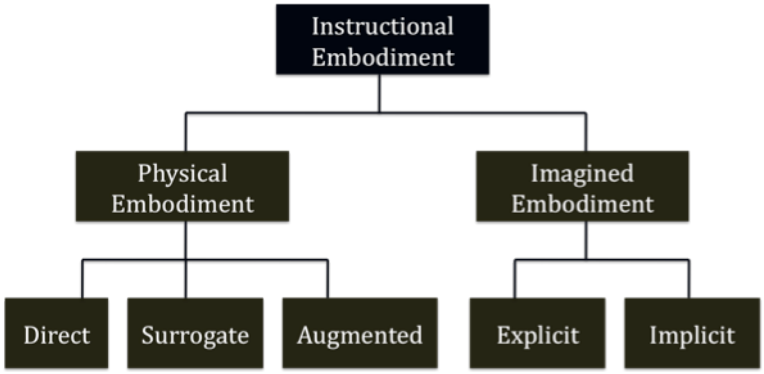


Figure 1. Instructional Embodiment Framework (Black et al., 2012).

The study presented in this paper investigated three different types of embodiment identified by the *Instructional Embodiment Framework* (i.e., Direct Embodiment, Surrogate Embodiment, and Explicit Imagined Embodiment), and compared these three types of embodiment with a no embodiment control condition. It may be noted that the group Imagined Embodiment in this study refers to Explicit Imagined Embodiment. Based on embodied cognition theory and previous research, I hypothesized that both physical and imagined embodiment would enhance learning and motivation more than no embodiment. Since physical embodiment can provide learners with more opportunities to construct multimodal mental perceptual simulations than imagined embodiment, I also hypothesized that physical embodiment would enhance learning and motivation more than imagined embodiment. I was also interested in finding out how surrogate embodiment (in which the learner sits in front of a computer and uses an avatar in a virtual environment to perform actions) would compare with direct embodiment (in which a learner physically moves and performs actions). Li, Kang, Lu, Han, & Black

(2009) investigated the effects of surrogate embodiment and direct embodiment as teaching and learning methods on students' understanding of abstract programming concepts. They found that children who learned in a direct embodiment condition (in which they acted like robots) showed higher gains in terms of comprehension and engagement as compared to children in a surrogate embodiment condition (in which they controlled the movements of a teacher who acted as a robot). Based on these findings, I expected learners engaged in direct embodiment to score higher on learning and motivation measures than learners engaged in surrogate embodiment.

Method

Participants

Eighty-four (N = 84) adult graduate students from a university in the United States participated in the study for course credit. Participants' age ranged from 21 to 50 years. A large majority (90%) of the participants were 21-30 years old, 50% were 21-25 years old, and 25% were 26-30 years old. There were 60 females and 24 males. The participants were from diverse cultural backgrounds. Participants identified themselves as Asian (42%), Caucasian (39%), African American (6%), Hispanic (5%), South Asian (4%) and Other (4%). All participants were proficient in English.

Design

The study employed a between-subjects posttest-only control group design. Four groups were investigated: (a) Direct Embodiment (DE), (b) Surrogate Embodiment (SE), (c) Imagined Embodiment (IE), and (d) a No Embodiment (NE) control. The dependent variables were memory retrieval, comprehension, near transfer, far transfer, and motivation.

Materials

The materials for the main learning task for all four groups included: (a) Novel historical text from the Indian subcontinent printed on paper about *Humayun*, India's Mughal Emperor, (b) illustrations and pictures of the main characters in the text, (c) Apple computers for the Surrogate Embodiment group, and (d) a blank A-4 size paper and pencil for the Imagined Embodiment group. Measurement materials included: (a) A memory retrieval paper and pencil test with twenty multiple-choice and open-ended questions that tested immediate recall of facts, (b) a comprehension paper and pencil test with ten open-ended questions that required participants to make inferences and think beyond the text, (c) a near transfer test, which was a history comprehension test, containing text from Indian history about the Queen of Jhansi that included elements common with the original text given during the main learning task followed by seven open-ended comprehension questions, (d) a far transfer test, which was a literature comprehension test, containing text from *Birbal*'s stories from Indian literature that included elements common with the original text given during the main learning task followed by seven open-ended comprehension questions, and (e) five motivation items on a questionnaire, which measured participants' enjoyment, confidence in their learning, increased interest in history, general interest and overall motivation. Participants responded to the motivation items on a five-point likert scale that ranged from strongly agree to strongly disagree. The maximum score for memory retrieval was 30, the maximum score for comprehension was 20, the maximum score for the near transfer test was 25, the maximum score for the far transfer test was 25, and the maximum score for motivation was 25.

Procedure

After the informed consent process, participants were randomly assigned to the four groups without knowing what condition they were in. All participants were given the same printed text, which they read silently one time, and they viewed pictures of the main characters in the text (these were illustrations and pictures of avatars). Next, participants were given the main learning task. The total time allocated to complete the task was 15 minutes. Participants in the Surrogate Embodiment group were given a brief tutorial and practice session before beginning the main learning task to familiarize them with their avatars and the features they were to use in a multi-user virtual environment. All participants were instructed to reread the text during the learning task.

For the learning task, participants in the Direct Embodiment group were instructed to physically play the role of the main character in the text (see Figure 2). All participants in the group interacted with and role-played with the experimenter, who played the role of another important character in the text. Participants in the Surrogate Embodiment group were instructed to play the role of the main character through an avatar (or virtual agent) and engage in virtual role-play in a multi-user environment (see Figure 3). Participants used virtual gestures and movements during the role-play. A confederate in a remote location controlled an avatar that represented the other character used in the role-play. The experimenter used the same script and the same characters in both role-plays. Based on the script, participants were free to create their own dialogue during the role-play. The Imagined Embodiment group was instructed to reread the text imagining the characters and the

actions in the text. The No Embodiment (control) was not given any instructions other than to reread the text silently to control for time.



Figure 2. Physical Role-Play During Direct Embodiment.



Figure 3. Avatar Role-Play During Surrogate Embodiment.

After the learning task, all participants were given the memory retrieval test, comprehension test, near transfer test, and far transfer test. Participants were given a maximum of 15 minutes to complete each test. To create a delay between immediate recall and the transfer tests, a distraction task was used before the near transfer test in which the experimenter had a brief conversation with the participants. The experimenter followed the same script for all distraction conversations. After the far transfer test, all participants were asked to complete a questionnaire containing motivation items. The questionnaire also contained questions about participants' backgrounds, their attitudes and opinions about technology and history, and manipulation checks. All items other than the background questions required participants to select responses on a five-point likert scale that ranged from strongly agree to strongly disagree. The questionnaire was followed by a feedback session in which participants were asked questions to gain an insight into how they learned. All participants were debriefed at the end.

Results

Multivariate tests results were found to be statistically significant at the .05 alpha level, Wilks' $\Lambda = .252$, $F(5, 76) = 9.089$, $p < .001$, $\eta 2 = .369$. Groups were found to differ significantly on: (a) memory retrieval, $F(3, 80) = 21.543$, $p < .001$, $\eta 2 = .447$; (b) comprehension, $F(3, 80) = 17.267$, $p < .001$, $\eta 2 = .393$; (c) near transfer, $F(3, 80) = 36.497$, $p < .001$, $\eta 2 = .578$; (d) far transfer, $F(3, 80) = 40.167$, $p < .001$, $\eta 2 = .601$; and (e) motivation, $F(3, 80) = 10.287$, $p < .001$, $\eta 2 = .278$.

Table 1: Mean scores and standard deviations for all dependent variables.

	Direct Embodiment Mean (*SD*)	Surrogate Embodiment Mean (*SD*)	Imagined Embodiment Mean (*SD*)	No Embodiment Mean (*SD*)
Memory Retrieval	23.95 (2.31)	22.38 (2.91)	18.81 (2.77)	18.12 (3.03)
Comprehension	16.09 (2.45)	16.71 (1.65)	14.09 (2.93)	11.67 (2.82)
Near Transfer	16.86 (2.85)	15.28 (2.74)	10.52 (2.94)	9.14 (2.69)
Far Transfer	15.33 (3.15)	12.86 (2.97)	8.38 (2.11)	7.57 (2.29)
Motivation	22.14 (3.21)	20.76 (2.99)	19.24 (2.07)	17.52 (2.94)

Post-hoc Tukey HSD tests revealed that overall Direct Embodiment and Surrogate Embodiment groups performed better than Imagined Embodiment and No Embodiment groups. Direct Embodiment and Surrogate Embodiment did not differ significantly from each other on memory, comprehension, near transfer, and motivation, $p > .05$, but they did differ significantly on far transfer. Direct Embodiment scored significantly higher on far transfer than Surrogate Embodiment, $p = .018$. See Table 1 and Figure 4.

Both Direct Embodiment and Surrogate Embodiment scored significantly higher than Imagined Embodiment and No Embodiment on memory retrieval, near transfer, and far transfer, $p < .001$. Direct Embodiment scored significantly higher than both Imagined Embodiment ($p = .007$) and No Embodiment ($p < .001$) also on motivation. Surrogate Embodiment scored significantly higher than No Embodiment ($p = .002$) on motivation but it did not score significantly higher than Imagined Embodiment on motivation, $p > .05$. Surrogate Embodiment scored significantly higher than both Imagined Embodiment ($p = .006$) and No Embodiment ($p < .001$) on comprehension. Although Direct Embodiment scored significantly higher than No Embodiment ($p < .001$) on comprehension, it did not score significantly higher than Imagined Embodiment on comprehension, $p > .05$. This suggests that although Direct Embodiment and Surrogate Embodiment did not differ significantly on comprehension and motivation, Direct Embodiment enhanced motivation more than Imagined Embodiment while Surrogate Embodiment did not enhance motivation more than Imagined Embodiment. In contrast, Direct Embodiment did not significantly improve comprehension more than Imagined Embodiment while Surrogate Embodiment was found to improve comprehension more than Imagined Embodiment. See Table 1 and Figure 4.

Another interesting finding is that participants in the Imagined Embodiment group scored significantly higher than participants in the No Embodiment group on comprehension ($p = .013$), but the two groups did not differ significantly on any other dependent variable, $p > .05$. See Table 1 and Figure 4.

No statistically significant relationship was found between participants' gender, age, ethnicity, cultural background, and attitudes towards technology and history and the dependent variables.

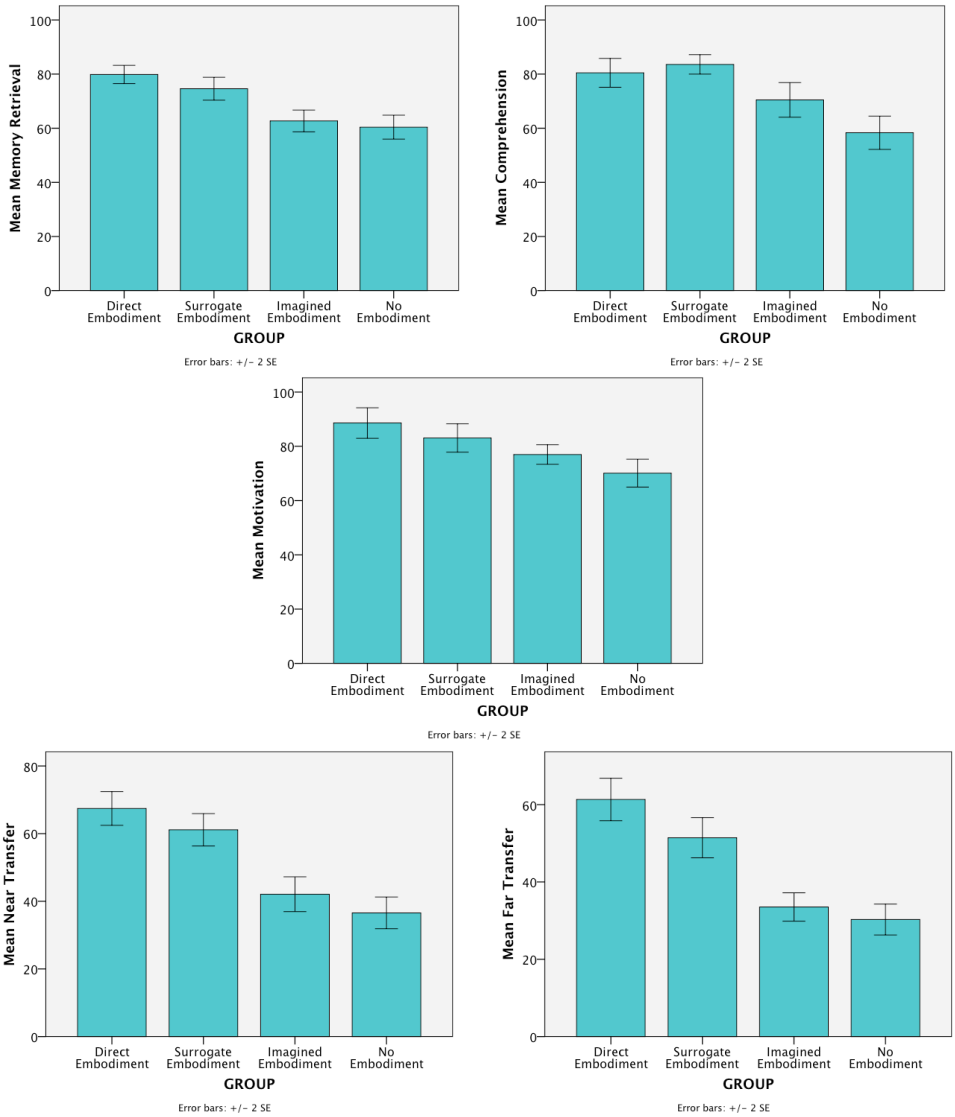


Figure 4. Mean Percentages for Memory Retrieval, Comprehension, Near Transfer, Far Transfer, and Motivation for All Groups.

Discussion

The results suggest that embodiment enhances learning and motivation and that constructing imaginary worlds improves comprehension. Our first hypothesis was that all three types of embodiment would enhance learning and motivation more than no embodiment. I found that participants who experienced both types of Physical Embodiment (i.e., Direct and Surrogate) scored higher than No Embodiment on all dependent variables. This supports previous research on embodied learning and instruction. I also found that similar to Direct Embodiment and Surrogate Embodiment, Imagined Embodiment enhanced comprehension more than No Embodiment. This finding supports research by Glenberg et al. (2004) who found that imagined manipulation enhances comprehension more than a no-manipulation read reread control condition.

Although I also hypothesized that Imagined Embodiment would score higher than No Embodiment on all dependent variables, I found that Imagined Embodiment did not enhance memory, near transfer, far transfer and motivation more than No Embodiment. This does not support previous research findings that suggest that imagining actions enhances memory and motivation more than no embodiment. For example, Glenberg et al. (2004) found that imagined manipulation enhanced memory and transfer more than a control condition. One reason for Glenberg et al.'s (2004) significant results could be that their participants were children who were given a relatively simple task. In our study, adult participants were given a completely novel text about the history of a country they were not very familiar with. The lack of a significant difference between Imagined Embodiment and No Embodiment scores could be attributed to the level of difficulty and the novelty of the text. Another possibility, one might argue, is that although the No Embodiment group was not explicitly instructed to imagine actions, participants in this group were engaged in implicit imagination. Since Black et al.'s framework

especially considers Implicit Imagined Embodiment, I included manipulation checks that gathered information about whether or not participants in the No Embodiment group imagined actions while reading. Although participants reported that they did not imagine actions while reading, we must acknowledge that these were self-report measures.

I also hypothesized that Physical Embodiment (i.e., Direct and Surrogate) would enhance learning and motivation more than Imagined Embodiment. I found that participants in the Direct Embodiment group scored higher than participants in the Imagined Embodiment group on all dependent variables except for comprehension. Participants in the Surrogate Embodiment group scored higher than participants in the Imagined Embodiment group on all dependent variables except for motivation. The lack of significant gains in comprehension for Direct Embodiment compared with Imagined Embodiment could be attributed to attention. It may be assumed that since participants had to act physically and move around physically, this distracted them. On the other hand, one might argue that since Direct Embodiment and Surrogate Embodiment comprehension scores did not differ significantly from each other, attention might not be the issue. In that case, it seems more plausible that imagination plays a significant role in helping people understand text and our results indicate that imagining actions improves comprehension as much as physically performing actions. This also supports previous neuroimaging and empirical findings.

Direct Embodiment and Surrogate Embodiment groups did not differ significantly on memory, comprehension, near transfer and motivation, which suggests that virtual role-play was just as effective as physical role-play. Participants in both groups reported that they were able to imagine themselves as the main character during role-play and that helped them relate to the reading. Their feedback suggests that physical and virtual role-play made learning equally meaningful for them. These results do not support Li et al.'s (2009) findings that Direct Embodiment enhances comprehension more than Surrogate Embodiment. However, we must keep in mind that Li et al. investigated a different domain.

One interesting finding is that Direct Embodiment far transfer scores were significantly higher than Surrogate Embodiment far transfer scores. This suggests that physical role-play might have an advantage over virtual role-play with regards to far transfer. The question is why did participants in the Direct Embodiment group score significantly higher on far transfer than participants in the Surrogate Embodiment group when the two groups did not differ significantly on any other dependent variable? Direct Embodiment possibly involved the reactivation of more multimodal representations during simulation, and this might have contributed to the far transfer. Feedback from participants also revealed that most participants in the Direct Embodiment group were able to imagine themselves as the main character in the text provided in the far transfer test. They also reported enjoying this text more than participants in the Surrogate Embodiment group because they were able to imagine the story and imagine themselves as the main character. In this sense they seemed to be transferring the imaginary worlds construction from the main learning task to the far transfer test.

Conclusion and Implications

The findings suggest that embodied instruction via role-play can make learning more meaningful and it can significantly enhance learning, transfer, and motivation more than no embodiment. The findings also suggest that virtual role-play via avatars can be as effective as physical role-play in enhancing memory, comprehension, near transfer, and motivation. This has implications for teaching and learning. Physical role-play might not be feasible in classrooms due to time constraints. Physical role-play is also not possible in distance learning. The study, therefore, provides evidence for using surrogate embodiment via avatar role-play for teaching and learning. The findings also support theory and research on the role of imagination in learning. This suggests that embodiment can enable learners to approach learning in a manner that is enjoyable and relevant to them, which makes learning more meaningful.

References

Barab, S., Dodge, T., Thomas, M. K., Jackson, C., & Tuzun, H. (2007). Our designs and the social agendas they carry. *The Journal of the Learning Sciences*, *16*(*2*), 263-305.

Barsalou, L. W. (2008a). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

Barsalou, L. W. (2008b). Grounding symbolic operations in the brain's modal systems. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 9-42). Cambridge: Cambridge University Press.

Bianchi-Berthouze, N., Kim, W. W., & Patel, D. (2007). Does body movement engage you more in digital game play? And why? *Lecture Notes In Computer Science*, *4738*, 102-113.

Black, J. B. (2007). Imaginary worlds. In M. A. Gluck, J. R. Anderson, & S. M. Kosslyn (Eds.), *Memory and mind* (pp. 195-208). NJ: Lawrence Erlbaum Associates.

Black, J. B., Khan, S. A., & Huang, S. C. D. (2014). Video games as grounding experiences for learning. In F. C. Blumberg (Ed.), *Learning by playing: Frontiers of videogaming in education* (pp. 290-301). New York: Oxford University Press.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. L. (2012). Embodied cognition and learning environment design. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (pp. 198-223). NY: Routledge.

Boroditsky, L., & Prinz, J. (2008). What thoughts are made of. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 98-115). Cambridge: Cambridge University Press.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400-404.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355-263.

Gibbs, R. W. J. (2007). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. NY: Houghton Mifflin.

Glenberg, A. M. (2008). Toward the integration of bodily states, language, and action. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 43-70). Cambridge: Cambridge University Press.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*, 424-436.

Hauk, O., & Pulvermuller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping*, *21*, 191-201.

Khan, S. A. (2012). *Surrogate embodied learning in Second Life* (Doctoral Dissertation). Ann Arbor, MI: Teachers College, Columbia University, ProQuest, UMI Dissertations Publishing (3512509).

Khan, S. A. (2013). Enhancing learning and motivation through surrogate embodiment in MUVE-based online courses. In A. E. Walsh, S. Khan, R. Gilbert, & M. Gale (Eds.), *iED 2013 Proceedings* (pp. 30-41). Boston, MA: Immersive Education Initiative.

Khan, S. A., & Black, J. B. (2013). *Enhancing learning and motivation through positive embodied affect and surrogate embodiment in virtual environments*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California.

Khan, S. A., & Black, J. B. (in press). Surrogate embodied learning in MUVEs: Enhancing memory and motivation through embodiment. *Journal of Immersive Education*, *1* (1).

Lankoski, P, & Järvelä, S. (2012). An embodied cognition approach for understanding role-playing. *International Journal of Role-playing*, *3*, 18-32.

Li, D., Kang, S., Lu, C., Han, I., & Black, J. B. (2009). Case studies of developing programming skills via embodied experiences in an after-school LEGO robotics program for elementary school students. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 2209-2216). Chesapeake, VA: AACE.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42,* 445-452.

Lu, C. M, Kang, S., Huang, S. C., & Black, J. B. (2011). Building student understanding and interest in science through embodied experiences with LEGO robotics, *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2011* (2225-2232). Chesapeake, VA: AACE.

Metcalf, S. J., Dede, C., Grotzer, T., & Kamarainen, A. (2009). *EcoMUVE: Design of virtual environments to address science learning goals*. Paper presented at the American Educational Research Association Conference, San Diego, CA.

Metcalf, S. J., Kamarainen, A., Grotzer, T. A., & Dede, C. J. (2011). *Ecosystem science learning via multi-user virtual environments*. Paper presented at the American Educational Research Association National Conference, New Orleans, LA.

Noice, H., & Noice, T. (2001). Learning dialogue with and without movement. *Memory & Cognition, 29,* 820–827.

Pulvermuller, F. (2008). Brain embodiment of category-specific semantic memory circuits. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 71-97). Cambridge: Cambridge University Press.

Robbins, P., & Aydede, M. (2009). A short primer on situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 3-10). Cambridge: Cambridge University Press.

Scott, C. L., Harris, R. J., & Rothe, A. R. (2001). Embodied cognition through improvisation improves memory for a dramatic monologue. *Discourse Processes*, *31*, 293-305.

Shapiro, L. (2011). *Embodied cognition*. NY: Routledge.

Creativity as Practice(d) in a Design Studio

Christoph Richter, Julia Lembke, Elisa Ruhl and Heidrun Allert, Christian-Albrechts-Universität zu Kiel
Institute of Educational Science, Olshausenstr. 75, 24098 Kiel, Germany
Email: {richter, j.lembke, ruhl, allert}@paedagogik.uni-kiel.de

Abstract: Although the social and cultural dimension of creativity has been emphasized for quite some time, there is neither a consensus on how creativity can be nurtured nor on what it is to become creative. Adopting a practice-oriented perspective, this paper reports on an ethnographic study in a studio-based course on Interface Design. Drawing on observations and students' narrative accounts of their working processes, the local design studio is portrayed as a well-attuned system of structural elements, patterns of interaction and epistemic assumptions. The findings reveal basic similarities but also significant differences with other studies on educational design studios. It is suggested that these differences are due to differences in the epistemic frames enacted.

Introduction

Creativity has become a vital and highly valued aspect of science, technology, the arts, as well as the everyday life (e.g. Craft, 2011). It has been argued that the increasing interest in creativity is due to a global shift towards a knowledge-based society and innovation driven economy (Sawyer, 2008). As a consequence the question on how to promote creativity is of interest for policy makers and curriculum developers alike (Sawyer, 2012). However, creativity is not just relevant to ensure economic growth but also to address urgent social and ecological problems and to enable individuals to actively cope with the volatile, provisional, and precarious life-worlds they find themselves in. Creativity in this sense is not just a skill to be used in predefined settings and aimed to increase performativity, but essentially "can challenge the *status quo*" (Craft, 2011, p. 28). Adopting the conference's theme, it is hence important not only to ask how creativity can be nurtured but also what it is to become creative.

This paper reports on an ethnographic study carried out in a one-semester course in the study program on Industrial Design at the Muthesius Academy of Fine Arts and Design in spring 2013. Adopting a practice-oriented perspective, in which creativity is "seen as a mode of human interaction with the world" (Beardon, Ehn & Malmborg, 2002), the goal of the study was to describe respective patterns of interaction enacted by the students and the teaching staff and to trace the underlying epistemic frame in a design studio setting. The setting was chosen because design education in general and the design studio in particular is supposed to be geared towards the cultivation of creativity and should therefore promote respective practices. However, due to the situated nature of practices (e.g. Schatzki, 2012), the intent of our analysis is not to unravel general principles of creativity but to shed light on the meshwork of practice enacted in a particular pedagogical setting. Drawing on (1) observations of the actual doings and sayings of the students and the teaching staff during the contact hours, (2) students' narrative accounts of their working process and procedures, as well as (3) the material arrangements and artifacts present and utilized, the analysis shows how students work on the horizon of their (and others) knowledge, bring in personal perspectives and make deliberate use of the opportunities they spot to respond to the design challenge given.

The contributions of this study to the learning sciences are threefold. First, the study provides an account of the patterns of interaction occurring in a genuine pedagogical setting, hence adding to the overall educational case base. Second, it backs up the assumption that professional practices draw on specific epistemic norms (cf. Shaffer, 2004) and therefore challenges paradigmatic orientations in education. Third, it provides ideas on how creative practice might be nurtured within design but also in other domains.

A Practice-Oriented Perspective On Creativity

The concept of creativity has been an object of concern for quite some time across various disciplines (e.g. Hennessey & Amabile, 2010). While the social and cultural dimension of creativity has been emphasized since the 1980th current conceptualizations of creativity often still build on a model of the social as "an external environment, a set of stimulations that facilitate or constrain the creative act" (Glăveanu, 2010, p. 85) instead of conceiving creativity as an inherently social process. Similarly, most of the current accounts also hardly account for the way creativity is mediated by material artifacts and environments (e.g. Vyas et al., 2009).

The conceptual framework for the present study is therefore based on a practice-oriented perspective on creativity. From this perspective creativity is neither a property of a person, process, product, nor environment, but a way of interacting with the world. More precisely, creative practices can be understood as those modes of interaction in which individuals or collectives aim to cope productively with an otherwise

indeterminate situation, i.e. a situation that is inherently disturbed, confused, ambiguous, or unsettled (cf. Miettinen, 2006).

Drawing on the work of Schatzki (2001, 2012) and Hörning (2001, 2004), we take practices as the central unit of analysis, which are understood as "embodied materially mediated arrays of human activity centrally organized around shared practical understanding" (Schatzki, 2001, p. 2). A practice-oriented perspective holds that human action is not a sequence of self-contained intentional acts realizing preconceived plans but an ongoing process in which the human actors actively frame, respond to, and transform the situations they find themselves in, making use of the artifacts and resources available. Competent action in this sense entails both an adaption of the individual to the environment s/he is acting in, giving rise to habits and routine ways of doings things, but also procedures that allow the actor to cope with those situations in which established habits do not work. Social practices, including creative ones, are neither an accumulation nor an abstraction of individual acts, but those patterns and styles of action that emerge from repeated interaction, allowing the participants to form shared expectations on how to act in a certain situation (cf. Hörning, 2001). Social practices hence can be understood as the conventions and arrangements enacted by a certain group of people at a certain point in time. As practices cannot be separated from the concrete doings and sayings of the practitioners and the material assemblages in which these take place, particular practices are necessarily local and historical. Accordingly there is no such thing as a creative practice, but a multitude of creative practices enacted in various settings. Furthermore, as situations are usually open to different interpretations, they require an active framing of those engaged in them. Enacting a social practice therefore requires not only practical knowledge, but also knowledge about the overarching schemes that allow actors to interpret and define the situation they find themselves in (cf. Hörning, 2004). The practical know-how as well as the interpretive schemes can be understood as repertoires the actors use to cope with the situations they are facing.

Learning to become creative hence is about learning to engage in the continually evolving process of a creative practice. From a practice-oriented perspective on learning "the practitioner is an embodied subject produced through participation in practices that shape skills, knowledge, understanding and disposition to action" (Hager, Lee & Reich, 2012, p. 7). Learning to be creative is a situated process entangled with the development of the learner's identity. In this process learners not only develop an understanding of the domain and practical know-how but also interpretive schemes to draw on when facing a certain situation. An important aspect of these schemes is what Shaffer (2006) has called the epistemic frame. Epistemic frames are "the ways of knowing, of deciding what is worth knowing, and of adding to the collective body of knowledge and understanding of a community of practice" (Shaffer, 2006, p. 223). Such frames might intersect and overlap with traditional disciplines but are essentially bound to local practices and continually transformed by their enactment. Learning therefore is not to be understood as a reproductive but as an inherently transactional process in which the learner as well as the context are evolving.

Against this background, the research agenda we pursue is not geared towards the identification of general principles of creativity but aimed to describe how creativity is practiced in a concrete educational setting. Trying to trace the underlying epistemic frame the present study aims to shed light on the utilization and creation of knowledge in creative design efforts as well as its impact on the practitioners.

Research on the Design Studio

The design studio has been variously identified as common denominator and essential constituent of design education across the disciplines (e.g. Brandt et al., 2008; Wang, 2010). The design studio is characterized by (a) open-ended projects the students work on over a prolonged period of time, (b) various types of structured review or feedback sessions focused on the evolving project work, and (c) a public presentation of the project outcomes (cf. Shaffer, 2003).

Even though some authors have depicted the design studio as a distinct and consistent pedagogical approach (e.g. Kuhn, 2001; Brandt et al., 2008) it has been argued that there are apparently significant differences regarding both content and methods in studio teaching between schools and even within departments (e.g. Ledewitz, 1985). In fact various alternative models for design studio teaching have been proposed building on substantially different assumptions on the nature of design and the role of the designer (e.g. Dutton, 1987; Ledewitz, 1985; Wang, 2010). Ledewitz (1985) already suggested that the practices actualized within the design studio depend on the stipulated model of design. In a more recent interview study Carvalho, Dong & Maton (2009) found that design disciplines not only differ with respect to the domain knowledge they deem relevant but also with regard to the epistemic assumptions they build upon, a fact not least reflected in respective educational efforts. Against this background it seems important to have a close look at the model of design and respective epistemic assumptions enacted in a specific context, before looking for commonalities that hold across settings in the first place.

While there is number of studies on various aspects of design studios in different domains (e.g. Lahti, Seitamaa-Hakkarainen, Hakkarainen, 2004; Maldonado et al., 2007; Vyas et al., 2009), only few studies explicitly aimed to elicit the epistemic frames enacted in these settings. Three noteworthy exceptions are the

ethnographic studies on design studios reported by Schön (1987), Shaffer (2003), and Sawyer (2012). In his analysis of design studios in architecture Schön (1987) has focused on the interactions between the student and the studio master and suggested these interactions to be essential to the development of a designerly way of thoughtful action centered around processes of framing, naming, moving, and evaluating. Shaffer (2003) followed students in an architectural design course at the MIT trying to elicit the structures underpinning the students' practices in the studio. To do so he depicted the way in which surface structures, pedagogy and the particular epistemology of the design process form a coherent system of activity. The epistemology he traces emphasizes the need for an individual interpretation of the design problem by the architect as well as the organization of the design process "around the development and articulation of expressive ideas" (p. 25). Finally, Sawyer (2012) carried out an ethnographic study at two professional schools of art and design aiming to identify general principles of the cultural model of the design studio. Synthesizing observation from interaction in the studio and interviews with instructors and students from a variety of domains, including among others interior design, illustration & architecture, he characterizes the teaching practices in the design studio as a form of "disciplined improvisation" (p. 34), in which the students are supposed to master a deliberate and effective design process. All three studies focus on how design is actually practiced within the design studio. However, both Schön and Sawyer, either implicitly or explicitly, assume that there is a common model of design. Only Shaffer's (2003) analysis fully accounts for the situatedness of the enacted epistemology, and therefore provides the most direct point of reference for the present study.

Research Design

The study was carried out in a design studio setting at the Muthesius Academy of Fine Arts and Design in spring 2013. The Muthesius Academy, founded in Kiel in 1907, is devoted to the systematic study of art and design. The school has about 500 students and offer bachelor and master programs in the fields of Fine Arts, Industrial Design, Communication Design, Spatial Strategies and Art Education. The course we followed was part of the study program on Industrial Design with a specialization on Interface Design. It was run by a professor and a research assistant. Eleven bachelor students in the 5th semester and six master students took part in the course that lasted from April to July, spanning a period of 14 weeks. Under the overall theme "simulation/simulator" the students were asked to define and carry out individual design projects. All students enrolled in the course were included in the study.

Our orientation towards practices led to a combined use of different research methods including (a) observations of the interactions between the students and the teaching staff during the contact hours, (b) students' narrative accounts of their working process, either voiced in students' interactions with the teaching staff or in informal interviews carried out by the research team, and (c) the material arrangements and artifacts present and utilized in the design studio. Data was recorded in the form of extensive field notes supplemented by photos and audio-recordings when feasible. All in all, a total of three observers conducted over 64 hours of site observations, taking part in over 90 individual feedback sessions as well as the students' final presentations. In parallel, the observers wrote memos following the sessions they attended and conducted a workshop with the students and the professor aimed to elaborate on the utilization of design artifacts in the middle of the term. Each of the observers has at least two years of teaching experience in a design related domain. Informed consent to take part in this study was obtained from all participants including the teaching staff.

To identify commonalities within the setting but also to trace variability, each of the students' projects has been treated as a distinct case in the analysis. Field and interview notes were organized into chronological case logs. Using an abductive approach, case logs and memos were used to surface patterns of interactions, which were then iteratively tested against the other cases until a stable set of patterns was found. Afterwards, the patterns and supporting data were used to trace the underlying epistemic frame.

Findings

In line with the analytic procedure, the presentation of findings starts with a general description of structural elements of the design studio. Against this background an overview of the actualized patterns of interaction is given and the underlying epistemic frame is reconstructed.

Structural Elements of the Design Studio

As pointed out by Shaffer (2003), the setup and organization of a design studio is strikingly different from a lecture hall, seminar room or classroom. The students in the course were provided with a large open workspace that they were only sharing with students of another course in the study program on Industrial Design. Within this workspace the students were free to setup permanent working areas, an option made use of by six of the students. The other students used the room as a temporal working and meeting space, especially during plenary meetings as well as the feedback session with the teaching staff. In addition the students had access to a variety of workshops on campus, an option used by the students working on hardware related projects.

While plenary meetings and feedback sessions with the professor took place during two fixed timeslots on Tuesdays and Fridays each week, the research assistant also dropped in the workspace in between. Apart from this, the students were free to decide when to work on their projects, each of them having access to the shared workspace 24 hours a day. With a calculated workload of 23 working hours per week the design projects where supposed to make up for most of the study time. In parallel, the students had to attend an introductory course on Human-Computer Interaction also given by the professor.

The overall project assignment "simulation/simulator" was introduced to the students by the professor right in the beginning of the semester. Besides a general motivation the students were asked to address the theme from a designerly perspective either by building on existing projects in other disciplines or taking a more artistic stance towards the question of simulation and reality. In the first plenary meeting the professor invited the students to reflect on previous project experiences followed by a collective brainstorming and discussion of potentially relevant questions. In the two subsequent meetings the students were asked to present their personal working plans for the course, detailing the envisaged stages of their projects and time management, elaborate on potential design questions and give short presentations on a range of topics approaching the overall project theme from different perspectives. From then on the focus shifted towards the students' individual projects only interrupted by a plenary session in the beginning of May in which the students were asked to reflect on their work process, as well as a plenary presentation of interim results asked for by two students in the end of May. Apart from these plenary sessions most of the contact time, about 7-8 hours a week, were spent on individual, sometimes also small group, feedback sessions. In these sessions, which took place in the students' workspace, the students presented their work in progress and discussed problems, design options and future directions with the teaching staff. The feedback sessions were of varying length but usually lasted for about 20 to 40 minutes.

The projects ended in a plenary presentation of the project results, attended by the participants of the course and the teaching staff, as well as a public exhibition on the campus of the Muthesius Academy. Additionally projects had to be documented on an online platform provided by the academy.

Patterns of Interaction

The analysis of the case logs resulted in the formulation of twelve patterns of interaction, which synthesize the observations throughout the 14 weeks of students' project work. These patterns describe recurrent ways of how the students and teaching staff coped with and transformed the situations they were facing throughout the design process. They are supposed to provide middle-level abstractions in that they capture situationally bound regularities in a form potentially verifiable and intelligible to other practitioners (cf. Dearden & Finlay, 2006). Even though not every pattern was observed in each case, the set of patterns is assumed to be characteristic for this context in that each of the patterns was instantiated in at least 50% of the projects, often repeatedly. According to their spatio-temporal extension, the patterns have been grouped into three main clusters: (1) foundational patterns that provide a background and reference point for all other design activities but also locate students' projects in broader realms of personal and professional development, (2) structuring patterns that render resources accessible and orchestrate project activities, and (3) patterns geared towards the advancement of project related ideas. The latter includes both prospecting patterns (3a) aimed at the exploration of ideas as well as anchoring patterns (3b) focused on the safeguarding and integration of ideas. In the following we briefly sketch the patterns of interaction along these clusters.

Foundational Patterns

Foundational patterns of interaction include the explicit *framing and re-framing of the design space*, the *presentation of results to the outerworld* as well as the *working at the horizon of one's own capabilities*. Being provided only with the generic theme "simulation/simulator" the students had not only to produce a product but also to advance a frame of reference that motivates their design by circumscribing the design space they want to operate in. While for example some of the students framed simulations as a means to learn certain concepts others conceptualized it as a tool to open up new perspectives or to provoke emotions and trigger thoughtful reflection. Successful framing and eventual re-framing provided the students with concepts and criteria to communicate, focus, and orient their design project. In taking a certain perspective the students also had to position themselves in relation to disciplinary questions as well as societal concerns. Similarly, by being asked to present their results to interested audiences, both during as well as at the end of the course, the expected quality and relevance of students' projects, which were expected not only to foster personal learning but also to bring forth worthwhile concepts and products, became salient. Furthermore, by defining projects at the horizon of the students' capabilities, students were asked to move beyond the already known and learned, and explore into emerging opportunities.

Structuring Patterns

This cluster of patterns includes *agile project management*, *help seeking,* and *carving space*. In the course of their projects the students had to plan and manage their activities taking into account given constraints as well as

all kinds of uncertainties. While the participants apparently drew on a generic model of the design process including ideation, conception, design, prototyping, and presentation, the actual project management was highly agile in that the students adapted and revised their plans in light of the obstacles but also possibilities that opened up in the course of the semesters. The students did not only utilize given resources but actively sought to render new resources accessible. In particular, students were seeking help not only among fellow students and teaching staff but also among friends, relatives, external practitioners, and domain experts. In doing so the students not only resolved acute problems but also broadened their own scope of action and expanded or consolidated their social networks. In addition, students carved both individual and collective spaces in support of their projects rendering accessible both social as well as material resources.

Prospecting Patterns

The prospecting patterns of interactions that are central to almost all projects include: *imaginative walkthroughs*, *making ideas tangible, playing with ideas* and *reflective prototyping*. What these patterns have in common is that they explore into the design space aiming at new insights regarding potential constraints or potentialities. Even though the patterns address somewhat different situations, they all entail a momentum of uncertainty and limited knowledge. In an imaginative walkthrough the actors simulated an anticipated usage scenario trying to develop an empathic understanding of the foreseen target population and their experiences. These walkthroughs helped to identify requirements but also to elicit potential implications of a certain design decision. While these walkthroughs had a strong narrative moment, students created and made use of tangible objects when trying to come to terms with experiential qualities and bodily experiences relevant to their projects. In playing with ideas, the participants typically started from a vague idea or incident, which was then explored in an open-ended, associative, and non-judgmental manner. In playing with ideas verbal comments were riddled with gestures and comments but also augmented with finds, artifacts, as well as ad hoc sketches. Finally, the students also developed prototypes to explore the feasibility as well as potential (side-)effects of design options.

Anchoring Patterns

The anchoring patterns of interaction that complement the prospecting patterns comprise of the *focused lead-in and lead*-out as well as the *deliberate decision-making*. The focused lead-in and lead-out brackets the stream of events marking the start and end of the feedback sessions as well as all types of presentations. While the focused lead-in aimed to raise the dialogue partners' interest, provided required background information and set the agenda, the focused lead-out synthesized the outcomes of the session, including the steps to follow. Deliberate decision making, in contrast, was triggered whenever students realized that they were approaching a relevant bifurcation point. Rather than striving for a satisficing option only, students usually explored and elaborated on a set of design options before coming to a defensible decision.

Reconstruction of the Epistemic Frame

While each of the patterns of interaction denotes an important transformation in the course of the students' design projects, the patterns do not exist in isolation but form a complex meshwork. For example, a successful *imaginative walkthrough* usually requires a *focused lead-in and lead-out* in which a certain *framing* is introduced or challenged. The question hence arises whether there is a common interpretative scheme i.e. an epistemic frame against which this meshwork of practice is enacted and can be understood. Following Stumpf and McDonnell (2001) we reconstruct the underlying epistemic frame along the model of the design task, the model of the design process, as well as the model of the designer implied in the meshwork of practice.

Model of the Design Task

Despite the considerable differences in the ways students carried out their projects, ranging from highly experimental to concept driven approaches, a concern essential to all projects has been the development and conveyance of a sound and appropriate interactional experience. While the overall theme „simulation/simulator" is open to a variety of interpretations the students were expected to develop a perspective through which they want to approach the design task. This was already made explicit in the initial meeting, when the professor explained that: „design is anything but arbitrary". Throughout the course the professor urged the students to take a stance and make deliberate decisions based on their interpretation of the design task. Irrespective of the particular perspective the design space was however approached holistically. Functional, technical, experiential, aesthetical, and ethical issues were not treated separately but approached in a highly integrative fashion. For example, envisioned interaction metaphors for a mobile app were discussed not only in terms of their usability and visual appearance but also with regards to their meaning for a community of users. As a consequence, students and teaching staff were constantly cross checking for example how technical and aesthetical decisions would affect the experiential or ethical qualities of the designs. From an epistemic point of view the students were hence supposed not only to develop a concept or prototype, but an understanding of the creation of an interactive product and its qualities of use, based on a viable yet value-laden perspective.

Due to a strong focus on the envisaged qualities of use, in contrast to more technology-centered approaches, the elicitation and communication of the intended users experience were another major concern across projects. A student put his emphasis on the qualities of use this way: „My aim is not that the prototype works in the first place, but that the feeling it gives is a good one.“ Being aware that many phenomena relevant to the quality of a product are only insufficiently captured in abstract representations the professor warned a student: „But you cannot say, concept, concept, concept and then eventually comes the design ... this is exactly the tricky point.” As a consequence prototypical realizations of the designs were often the only way to convey experiential qualities the students were interested in.

Furthermore, with its emphasis on the qualities of use of a specific product, the focus was on the particular rather than on the universal. Even though the professor repeatedly stated that design is not necessarily about innovation and the students are supposed to built on and integrate rather than invent technologies they were however expected to develop original and also novel solutions. However, originality and novelty in this conception are bound to the particular. Due to the specificities of their design concepts, existing know-how or expertise repeatedly rendered pointless, forcing the students to carry out practical experiments in order to deepen the understanding of given design options. In this sense the students were expected not only to work on the horizon of their own capabilities but also to add to the disciplinary knowledge base.

Model of the Design Process

All in all the activities in the design studio at the Muthesius Academy were organized around the development of meaningful/fruitful options within reach. Even though the students did not have to develop a fully functioning product, they were expected to devise a design that at least in principle could be implemented with existing technologies and/or provide a working prototype conveying essential qualities of use. Towards this end the students continuously framed the design space, explored into and decided on design options they deemed most promising. While the initial framing of the design space marked an essential milestone for all projects, it was constantly reassessed and concretized in the course of the design process, sometimes resulting in a fundamental redirection of the overall project. With the emphasis on the development of sound solutions, in line with the student's interpretation of the design task, the students were neither asked to adopt a particular process model nor were they given a fixed set of design principles or methods. Procedures and criteria were rather suggested and agreed upon on a needs basis taking into account the particularities of the project at hand.

Despite the purported linear organization of the design process, chaining up phases of ideation, conception, design, (prototypical) realization, and presentation, the underlying epistemic processes of framing, exploration and deliberate decision making were highly iterative and agile in the sense that participants continuously reflected on the implications of the design moves made. Rather than drawing on a fixed set of requirements and constraints for an envisaged product, the participants sought and created situations allowing them to probe their ideas and provide new information and insights throughout all stages of the design process. Respective strategies such *as imaginative walkthroughs*, *making ideas tangible*, *playing with ideas* and *reflective prototyping* all typify forms of non-monotonic reasoning and hence expand the knowledge base, the participants can draw upon. Or as the professor put it with regard to prototyping: “It is particularly important, that there is something that you can figure out.” In the same way the design artifacts created by students were used a catalysts for further elaborations rather than as mere explications of preexisting ideas.

Model of the Designer

In taking a certain perspective on the design task and devising a solution, that is publicly exhibited, the students were not only expected to demonstrate their competencies and skills but also to position themselves in relation to disciplinary questions as well as societal concerns. The students were also expected not to stay with the already known and learned but to grow with their projects and produce meaningful results. The designer in this setting was characterized as a capable creator and decision maker who is able to cope with uncertain, complex and value-laden situations. At the same time, the designer was also expected to be aware of the limits of his own knowledge and skills. This dual demand was also apparent in the professor's behavior. He, at various occasions, articulated the limits of his own know-how while also expressing personal preferences and convictions.

Even though the outcomes of the design process were largely unpredictable, the designers were supposed to actively seek and explore the opportunities that are opening up. In doing so, they recurrently had to share preliminary and half-baked ideas as well as to put their models, mock-ups, and prototypes to test. While entailing the risk of failure, disappointment or misunderstanding this was seen as an important move, or as a student put it: “The more feedback you get, the more impressions you get.” Additionally, a general curiosity and openness towards novel things and ideas seems to be required. In an interview a student explained: “In the end it's the job of the designer to deal with superficial knowledge. As a designer you might be provided with a short briefing and then you have to work with it […] therefore its good that we are introduced to so many different subject matters.” Furthermore, asking for assistance and help was not only seen as legitimate but actively promoted by the teaching staff.

Discussion

Our findings fit quite well the overall characterization of the design studio as "a vital complex of material representation, social collaboration, creativity, emotionality and a tolerance for uncertainty if not outright confusion – balanced with a faith that meaningful designs eventually will emerge" Wang (2010, p. 176). On a general level the outcomes of our analysis also appear to be compatible with findings of Schön (1987), Shaffer (2003), and Sawyer (2012) in that (a) design is aimed at unique and open-ended problems, which have to be framed by the designer, (b) design is an iterative process in which a series of intermediate design products is created and reflected upon, (c) that this process is mediated by generative feedback and social scaffolds, and (d) that design is best taught in the process of designing itself.

However, when having a closer look at our findings there also appear some noteworthy differences in the way the design studio is enacted in the course we followed. First, in comparison to the other studies, the design assignment in the design studio at the Muthesius Academy was much more open, inviting students to build on existing projects in various disciplines or to take a more artistic stance towards the question of simulation and reality. Hence, the students had to start from their own themes and ideas and argue for their relevance. Second, the professor neither provided the students with a consecutive series of assignments, as reported by Shaffer (2003), nor did he advocate any particular process model, as suggested by (Sawyer, 2012). Even though the participants occasionally referred to a generic model of the design process, the actual approaches differed significantly and were highly agile. The design approach enacted by the students, significantly differed from the analysis-synthesis model referred to by Sawyer (2012), in that the students actively sought and created situations to generate new information and insights. Third, while the creation of tangible products was also a major concern in the studios observed by Schön (1987), Shaffer (2003), and Sawyer (2012), we found a strong emphasis on the experiential qualities than their formal or representational properties. To test and convey their ideas the students in our case had to create first hand experiences rather than representations of the intended products. In that sense they not only had to express their ideas but also to create an (experiential) proof of concept. Finally, the focus in the design studio we followed has not only been on the mastery of disciplinary skills but also on the cultivation of personal design identities as well as the advancement of the disciplinary knowledge base. In fact, the disciplinary boundaries of the design studio were rather open, which also reflect the interdisciplinary roots of the field of Interface Design.

The case study design does not allow for generalizations to other contexts and the reported deviations might at least partly be attributed to disciplinary differences or personal attitudes and preferences of the teaching staff. Additionally, as epistemic frames are ephemeral in nature they are not open to direct observation but must be inferred. However, despite these limitations the results challenge the assumption that the design studio builds on a uniform pedagogy and entails a particular epistemology. The findings rather indicate that we should expect substantial differences in the way the design task, the design process, and the designer are understood by those involved in respective practices.

Summary

The study traced the creative practices enacted in a design studio in the field of Interface Design throughout a semester. In the analysis the design studio was portrayed as a well-attuned system of structural elements, patterns of interaction and epistemic assumptions. The comparison of our findings with other ethnographic studies on educational design studios revealed some basic similarities, but also a range of significant differences. We argue that these differences are not incidental but back up the assumption that there are significant differences in the epistemic frames enacted by practicing designers as well as in design education. The respective assumptions about worthwhile forms of knowledge, forms of knowing, and the means to advance the collective knowledge base, have direct implications for the understanding of creative practice as well as what it means to become a creative actor.

The perspective taken in this study also raises to question the idea of a uniform design mode of thinking constitutive for all kinds of creative knowledge work as suggested for example by Bereiter (2010). Instead of striving for generic principles on how to foster creativity (e.g. Sawyer, 2012) we believe it to be more fruitful to continue the detailed analysis on how creativity is practiced in different domains and settings and shed light on the mechanisms through which respective practices are nurtured and cultivated.

References

Beardon, C., Ehn, P., & Malmborg, L. (2002). Design of augmented creative environments. *Proceedings of CSCL 2002, Boulder, Colorado, January 7-11, 2002,* Hillsdale: Lawrence Erlbaum Associates, pp. 503-504.

Bereiter, C. (2010). Where the Learning and Pedagogical Sciences Need Philosophers. In *Encyclopaedia of Philosophy of Education*. Retrieved 1 November, 2013 from http://www.ffst.hr/ENCYCLOPAEDIA/.

Brandt, C., Cennamo, K., Douglas, S., McGrath, M., Reimer, Y., & Vernon, M. (2008). (de)Coding the Studio Method to Teach the Design of Human-Computer Interaction. In *Proceedings of the 24th National Conference on the Beginning Design Student, Atlanta, Georgia, March 2008,* pp. 0-11.

Carvalho, L., Dong, A., & Maton, K. (2009). Legitimating design. A sociology of knowledge account of the field. *Design Studies, 30*(5), 483-502.

Craft, A. (2011). *Creativity and Education Futures – Learning in a Digital Age*. Stoke on Trent: Trentham Books.

Dearden, A., & Finlay, J. (2006). Pattern languages in HCI: A critical review. *Human-Computer Interaction, 21*, 49-102.

Dutton, T.A. (1987). Design and studio pedagogy. *Journal of Architectural Education, 41*(1), 16–25.

Glăveanu, V.P. (2010). Paradigms in the study of creativity: Introducing the perspective of cultural psychology. *New Ideas in Psychology, 28*, 79-93.

Hager, P., Lee, A., Reich, A. (2012b). Problematizing Practice, Reconceptualising Learning and Imagining Change. In: P. Hager, A. Lee, & A. Reich (Eds.). *Practice, Learning and Change (pp. 1-14)*. Springer: Dordrecht.

Hennessey, B.A., & Amabile, T.M. (2010). Creativity. *Annual Review of Psychology, 61*, 569-698.

Hörning, K.H. (2004). Soziale Praxis zwischen Beharrung und Neuschöpfung. Ein Erkenntnis- und Theorieproblem. In: K.H. Hörning, & J. Reuter (Eds.). *Doing Culture: Neue Positionen zum Verhältnis von Kultur und sozialer Praxis (pp.19-39)*. Bielefeld: transcript Verlag.

Hörning, K.H. (2001). *Experten des Alltags*. Weilerswist: Velbrück.

Kuhn, S. (2001). Learning from the architecture studio: Implications for project-based pedagogy. *International Journal of Engineering Education, 17*(4/5), 349-352.

Lahti, H., Seitamaa-Hakkarainen, P., & Hakkarainen, K. (2004). Collaboration patterns in computer supported collaborative designing. *Design Studies, 25*(4), 351-371.

Ledewitz, S. (1985). Models of design in studio teaching. *Journal of Architectural Education, 38*(2), 2–8.

Maldonado, H., Lee, B., Klemmer, S.R., & Pea, R.D. (2007). Patterns of Collaboration in Design Courses: Team dynamics affect technology appropriation, artifact creation, and course performance. *Proceedings of CSCL 2007, New Brunswick, NJ, July 16-21, 2007,* International Society of the Learning Sciences, pp. 486-495.

Miettinen, R. (2006). The Sources of Novelty: A Cultural and Systemic View of Distributed Creativity. *Creativity and Innovation Management, 15*(2), 173-181.

Sawyer, R.K. (2008). Optimizing learning: Implications of learning sciences research. In: OECD (ed.). *Innovating to learn, learning to innovate (pp. 45-65)*. Paris: OECD.

Sawyer, R.K. (2012). Learning how to create: Toward a learning science of art an design. *Proceedings of ICLS 2012, Sydney, Australia, July 2-6, 2012,* International Society of the Learning Sciences, pp. 33-39.

Schatzki, T. R. (2001). 'Introduction. Practice Theory'. In: T. R. Schatzki, K. Knorr-Cetina, & E. von Savigny (eds.). *The Practice Turn in Contemporary Theory (pp. 1–14)*. London: Routledge.

Schatzki, T.R. (2012). A Primer on Practices. In: J. Higgs et al. (eds). *Practice-Based Education: Perspectives and Strategies (pp. 13-26)*. Rotterdam: Sense Publishers.

Shaffer, D.W. (2003). *Portrait of the Oxford design studio: An ethnography of design pedagogy* (Working paper Np. 2003-11). Madison, WI: Wisconsin Center for Education Research.

Shaffer, D.W. (2004). Pedagogical Praxis: The Professions as Models for Postindustrial Education. *Teachers College Record, 106*(7), 1401-1421.

Shaffer, D.W. (2006). Epistemic frames for epistemic games. *Computers & Education, 46*, 223-234.

Schön, D.A. (1987). *Educating the Reflective Practitioner*. San Francisco: Jossey Bass.

Stumpf, S.C., & McDonnell, J.T. (2001). Talking about team framing: using argumentation to analyse and support experiential learning in early design episodes. *Design Studies, 23*(1), 5-23.

Vyas, D., Heylen, D., Nijholt, A., & Van Der Veer, G. (2009). Collaborative practices that support creativity in design. *Proceedings of ECSCW'09, Vienna, Austria, September 7-11, 2009*, Springer: London, pp. 151-170.

Wang, T. (2010). A New Paradigm for Design Studio Education. *The International Journal of Art & Design Education, 29*(2), 173-183.

Acknowledgements

We would like to thank the students and the teaching staff at the Muthesius Academy of Fine Arts and Design in Kiel who made this study possible. The research leading to these results has received funding from the European Union's Seventh Framework Programme FP7/2007-2011 under grant agreement n° 318552.

Fostering Scientific Reasoning: A Meta-analysis on Intervention Studies

Katharina Engelmann, Frank Fischer, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München, Germany
k.f.engelmann@psy.lmu.de, frank.fischer@psy.lmu.de

Abstract: Pedagogical intervention for scientific reasoning is a highly relevant topic in science education and outside the classroom. A systematic analysis of the success of interventions on scientific reasoning is still missing, leaving unanswered questions regarding the magnitude of the effect of interventions for scientific reasoning; and which factors in the intervention and the assessment explain differences between studies. Effect sizes taken from 15 studies were included in a meta-analysis. The results revealed a large effect of interventions on scientific reasoning (g = 0.80). Moderator analyses included learning activities and, surprisingly, showed that constructive activities yielded larger effects than interactive ones. The meta-analysis is limited by the number of studies included. Nevertheless, the results show that scientific reasoning can be fostered, though the success of the intervention depends on variables in its content, pedagogy, and assessment.

Objectives and Purpose

Scientific reasoning has become a prioritized topic in science education. Science in school aims to teach more that scientific knowledge; it aims to develop a scientific way of reasoning (Zimmerman, 2000). However, scientific reasoning does not only play a role in science education. A lot of information outside the classroom has been generated in scientific research. It is possible to process information derived from science without being able to reason scientifically. Nevertheless, an understanding of the processes and concepts of science; the skill to apply this knowledge; and the ability to reason scientifically enables people to understand scientific information within its context, including the assumptions and limitations that derive from its origin (Giere, 1979). Fostering scientific reasoning is therefore not only relevant for science education; it is also relevant for enabling people to participate in a world in which everyone is surrounded by science-based information.

Zimmerman (2007, p. 215) states that the "…issue of the best way to assess the effectiveness of instructional interventions [for scientific reasoning] will be the next issue in need of resolution". This meta-analysis presents a first approach to an exploration of the effects of interventions on scientific reasoning. Furthermore, it analyzes to what extent factors in the intervention and in the measurement of scientific reasoning can be identified to moderate the effect of the interventions.

Theoretical Framework

Conceptualizations of Scientific Reasoning

There are several different conceptions of scientific reasoning; however, as diSessa noted, any effort toward a closed set definition of scientific reasoning is an "…elusive and likely chimerical goal" (diSessa, 2008, p. 560). Heuristically, conceptualizations of scientific reasoning can be differentiated into approaches that focus on scientific reasoning as a process in scientific inquiry (e.g. Klahr and Dunbar, 1988; Lawson, 1995); approaches that focus on scientific argumentation (e.g. Kuhn, 1993, 2010); and approaches that focus on an understanding of the principles of science (Giere, 1979).

One of the most frequently cited approaches to scientific reasoning (according to *Google scholar*) is the conceptualization by Klahr and Dunbar (1988) who understand scientific reasoning to be a scientific discovery that is conducted through a dual search process in an hypothesis space and an experiment space. The model includes three main components: the *search in the hypothesis space*, during which an hypothesis is evoked by prior knowledge or induced by observations from experiments; *test hypothesis*, during which experimentation is used to evaluate a specific hypothesis; and *evaluate evidence*, during which it is analyzed whether all the results that are produced regarding an hypothesis allow its rejection or acceptance of the hypothesis. In addition, further models of scientific reasoning were developed that focus on procedural aspects e.g. in science education (Lawson, 1995).

Argumentation is considered to be "…common in science" (Osborne, 2010, p. 463) and also serves as a pedagogical method to facilitate scientific reasoning (Osborne, 2010). Kuhn (1993) connected scientific reasoning to reasoning as an argument (1991). She understands argumentation to be either rhetorical or dialogic. A dialogic argumentation contains a minimum of two people with different views, who join a dialog in which everyone offers justifications for their own view and counterarguments opposing the other view. A rhetorical argument contains a juxtaposition of two opposite claims and a reasoning process in which the truth or falsity of

the claims is considered. Both forms of argument share a similar form of argumentative reasoning: at least one person notices the opposition between different claims and evidence is used in the following dialog or reasoning process to support or challenge claims. Each of the individuals involved in the dialog or reasoning process stays open with regard to the possibility that the original claim might be wrong, while allowing new evidence to have an impact on the evaluation of the assertion without dominating the reasoning process. In most cases, it appears that claims are not completely correct. In an ideal argument the evidence would be therefore "…weighted in an integrative evaluation" (Kuhn, 1991, p. 12) towards an integrative resolution. Therefore, scientific argumentation could be conceptualized as a rhetorical or dialogic argument about science, scientific constructs, and/or within a scientific context.

The nature of science contains a number of principles that are characteristic of the process of science. The American Association for the Advancement of Science (1989) understands the nature of science to be a particular way of observing, experimenting, validating, and thinking. Knowledge about the nature of science contains an understanding of scientific methods, the nature of scientific reasoning, and also a set of beliefs and attitudes about the world that serves as a foundation of science. Scientific reasoning, as an understanding and application of the nature of science, focuses on knowledge about scientific statements and their justification and arguments; the role of theories; statistical methods; the difference between causes and correlations; and values and decisions in science. While this description consists of mainly conceptual knowledge, understanding of the nature of science also includes the components of applying this knowledge within the scientific context; moreover, it also includes the application outside the scientific context in everyday situations (Giere, 1979). Here, the link between scientific literacy and scientific reasoning becomes apparent. Scientific literacy is considered to be an "…understanding of science concepts and processes with the assumption that such understanding would lead to an informed citizenry able to enact their knowledge in personal and societal issues" (Cavagnetto, 2010, p. 336).

The description of the differentiation between processes of scientific reasoning, scientific argumentation, and understanding of the nature of science includes a wide range of knowledge and skills. Therefore, it seems reasonable to adapt this differentiation to an analysis of differences between the contents of scientific reasoning in interventions that aim to foster scientific reasoning and test that measures the success of those interventions.

Fostering Scientific Reasoning

Empirical studies have begun to investigate the effect of interventions on scientific reasoning (e.g. Duncan & Arthurs, 2012); scientific inquiry (e.g. Gutwill & Allen, 2012); and scientific argumentation (e.g. Stark, Puhl, & Krause, 2009). Successful interventions were reported concerning children within a school context (e.g. Kuhn & Dean Jr., 2005), children outside a school context, e.g. in a museum (e.g. Gutwill & Allen, 2012), and also concerning adults in higher education (e.g. Duncan & Arthurs, 2012).

An analysis of pedagogical approaches to foster scientific reasoning can be conducted from the perspective of the intervention in which scientific reasoning is fostered and from the perspective of the assessment.

From the perspective of interventions, it could be hypothesized that factors that influence learning in general also influence the facilitation of scientific reasoning. The first factor is the differentiation in content of scientific reasoning as described in the conceptualization of scientific reasoning. Additionally, different types of knowledge could be targeted within each content of scientific reasoning. Moreover, Mayer (2012) suggests analyzing learning in terms of the knowledge type that is targeted in an intervention. He differentiates between *facts*, *concepts*, *processes*, *strategies*, and *beliefs*.

Apart from the content of the intervention, it can be assumed that pedagogical methods influence the success of an intervention. Chi (2009) suggests differentiating between the activities that learners undertake. She describes a framework (ICAP) distinguishing *active*, *constructive*, and *interactive activities*. *Active* activities include all activities during which the learner does something physically. *Constructive* activities require learners to produce something that goes beyond the information in the learning environment. *Interactive* activities can be differentiated into *instructional dialogs*, in which learner interacts with an expert or teacher, and peer dialogs, in which learners refer to each other and build on each other's contributions. In addition, it can be hypothesized that the technological support and the way in which the technology is supporting learners during the intervention influences the success of the intervention. Although the technological support became an important variable in the design of learning environments, a main conclusion from research on technology-enhanced learning is that the effects of technology are rarely main effects but rather interaction effects of technology with the pedagogical approach in which it is used. In recent years, many technology-supported interventions have been developed using a constructivist perspective by engaging learners in authentic activities, providing scaffolds for self-regulation and meta-cognition and encouraging collaboration (Rosen & Salomon, 2007).

From the perspective of assessment, the transfer distance that is demanded in the post-test could be understood to be represented by the degree to which the knowledge type included in the post-test was already

addressed in the intervention (e.g. Barnett and Ceci, 2002). It can be hypothesized that the transfer distance is negatively related to the size of the effect or, in other words: the effect of an intervention is larger if what is measured in the post-test is more similar to what has been facilitated during the intervention.

Research Questions

The scientific community – especially in developmental psychology and science education - has been interested in scientific reasoning and related constructs (Zimmerman, 2000). Moreover, there seems to be an interest in fostering scientific reasoning skills across different content areas. Even though several single studies yielded positive effects of interventions on scientific reasoning, a systematic analysis of the success of interventions for scientific reasoning and possible moderating factors is missing. Interventions use different pedagogical approaches and aim to facilitate different aspects of scientific reasoning. Interventions and post-tests differ across studies in the content of scientific reasoning, the knowledge type, the technological support, learning activities, and the degree to which the knowledge type included in the post-test was already addressed in the intervention; consequently, it could be hypothesized that these differences explain parts of the variability between the effects of interventions on scientific reasoning.

Research question 1: What is the magnitude of the effect of interventions on scientific reasoning? What is the variability of the effects across intervention studies?

Research question 2: If there is variability of effect sizes across studies, to what extent do the content of scientific reasoning, the knowledge type, the technological support, and the type of learning activities included *in the intervention* explain the variability of the effects between studies?

Research question 3: If there is variability across studies, to what extent do the content of scientific reasoning, the knowledge type, and the transfer with respect to different knowledge types *in the post-test* explain the variability between studies?

Method

Literature search and selection

A systematic literature search was conducted to identify and retrieve publications concerning the facilitation of scientific reasoning. The literature search was conducted in the databases PsyINFO and ERIC using the search terms *scientific reasoning*, *scientific thinking*, *scientific discovery*, *scientific inquiry*, and *scientific argumentation*, which resulted in 2722 papers. The search was conducted in March 2013. The results were restricted to literature that contained at least one of the five search terms in the title in order to reduce the number of findings and eliminate mostly irrelevant literature, which resulted in 664 studies.

The studies included in the meta-analysis were selected using the following criteria: (a) empirical publication in a scientific journal or book, (b) in English or German, (c) published within the period from 01.01.1988 through 31.12.2012, and inclusion of a report of (d) an intervention and (e) at least one between-group comparison in a post-test separate from the intervention. Some studies did not report all the necessary information needed to conduct a meta-analysis. In these cases, the authors were contacted via email. In case they did not respond, whenever possible the existing data were used if sufficient to enable inclusion of the publication in the meta-analysis, otherwise the publication was eliminated. The selection of literature resulted in 15 studies to be included in the meta-analysis.

Literature coding

Coding schemes were developed in order to analyze the content of scientific reasoning in interventions and post-tests; the knowledge type in interventions and post-tests; the technological support in interventions; and the learning activities in interventions.

An intervention was operationalized as a difference in treatment between an experimental and a control group. Three studies included more than one intervention. Abdullah and Shariff (2008) included two experimental groups, of which only one (the HACL condition) was included in the meta-analysis because the description of the study suggested that this condition was the author's target condition with respect to the intervention. Gutwill and Allen (2012) included two experimental groups, of which the Juicy Question group was chosen for the same reason. Zion, Michalsky, and Mevarech (2005) had three experimental groups, of which the most inclusive intervention condition was chosen (the condition that included the attributes of the two other experimental conditions). Furthermore, two studies had more than one control group. In both cases, the name of the control group led our decision on inclusion in the meta-analysis (i.e. "pure control" in Gutwill & Allen, 2012 and "main control" in Kuhn & Dean Jr., 2005).

Content of scientific reasoning: The coding scheme for the content of scientific reasoning was derived from the conceptualizations of scientific reasoning and differentiated between the *processes of scientific reasoning*, *scientific argumentation*, and *understanding science*. Intervention that fostered *processes of scientific reasoning* included scientific processes such as the deduction or generation of hypothesis, or the generation or

evaluation of evidence. Interventions that fostered *scientific argumentation* included an argumentation about science or scientific constructs. Interventions that fostered *understanding science* included knowledge about the principles, concepts, assumptions, and limitations of science and its application. One code was given to each intervention.

Knowledge type: The coding scheme for the knowledge type differentiated between *facts*, *concepts*, *processes*, *strategies*, and *beliefs* (Mayer, 2012) with respect to scientific reasoning. One or more codes were given to each intervention representing the knowledge types that were included in the intervention. In order to adequately analyze the data, the interventions were classified into (a) one group of interventions that included facts, concepts, or both (b) one group that included processes, strategies, or both and (c) a third group that included a combination of at least one item from each of the two prior groups.

Technological support: The coding scheme for the technological support differentiated between interventions that included technological support from a constructivist perspective, operationalized as creating an authentic learning environment, supporting cooperation, or supporting self-regulation; technological support that did not support any of these three aspects of a constructivist approach; and interventions that did not include any technological support. One code was given to each intervention.

Learning activities: The coding scheme for the learning activities (Chi, 2009) differentiated between *active*, *constructive*, interactive in an *instructional dialog*, and interactive in a *peer dialog*. One code was given to each intervention representing the most dominant activity.

All coding schema also included one residual category in case the description did not give enough information to determine which code was true. These data were treated as missing data in the moderator analysis.

The measurement of scientific reasoning was operationalized as the post-test(s) conducted after the pedagogical intervention. The coding scheme for *content of scientific reasoning* and *knowledge type* in the post-test was similar to the coding scheme for the intervention.

Transfer with respect to different knowledge types: The transfer was calculated by categorizing each study in respect to the distance between what was measured in the post-test to what was facilitated during the intervention. We differentiated between post-tests that included *no transfer* and posttest that included *transfer*. No transfer was assumed in cases where the post-test included only knowledge types that were already part of the intervention. Transfer was assumed in cases were the post-test also (or only) included knowledge types that were not part of the intervention.

Statistical analysis

The meta-analysis was conducted following the procedure suggested by Lipsey and Wilson (2001) for the fixed model. Wherever possible, the descriptive data were used to calculate the effect size (Hedges' g) by using the arithmetic means of experimental (X_G1) and control (X_G2) groups; the pooled standard deviation (σ_pooled); and the samples size for each group (n1 and n2): g=(XG1−XG2)/(σ_pooled), (σ_pooled)=√(σ12⁢(n1−1)+σ22 ⁢(n2−1))/(n1+n2−2).

Alternatively, The results of the inferential statistics (*t* or *F* values) were used to calculate the effect size using these formulas: g=t*√((n1+n2)/n1⁢n2) or g=√(F*(n1 + n2)/n1⁢n2).

Most studies reported one effect in respect to scientific reasoning and this effect was used in the meta-analysis. The arithmetic mean of the effect sizes was calculated for those studies that reported more than one outcome, except for one study (Duncan & Arthurs, 2012) who reported two outcome measures based on different but possibly not independent samples. In this case, the more reasonable effect size was chosen.

The specific parameters of the meta-analysis and the moderator analysis were calculated using the meta-analysis macros for SPSS from Wilson (2005). The effects were corrected for small sample bias and weighted using the inverse variance weight. The results were computed by calculating the mean of the effects and the confidence interval (CI). Furthermore, the homogeneity (Q) of the effects was tested.

Results

The first research question was: What is the magnitude of the effect of interventions on scientific reasoning? What is the variability of the effects across intervention studies?

The meta-analysis revealed a highly significant effect, suggesting a large mean effect size (g = 0.80, $CI_{95\%}$ [0.70, 0.90], p < 0.01). Figure 1 gives an overview of the effects included in the meta-analysis. The analysis of homogeneity was highly significant (Q (14) = 96.34, p < 0.01), showing that the sample is heterogeneous. The magnitude of the effects of interventions for scientific reasoning is high; however, the results show a high variability of the effects across intervention studies.

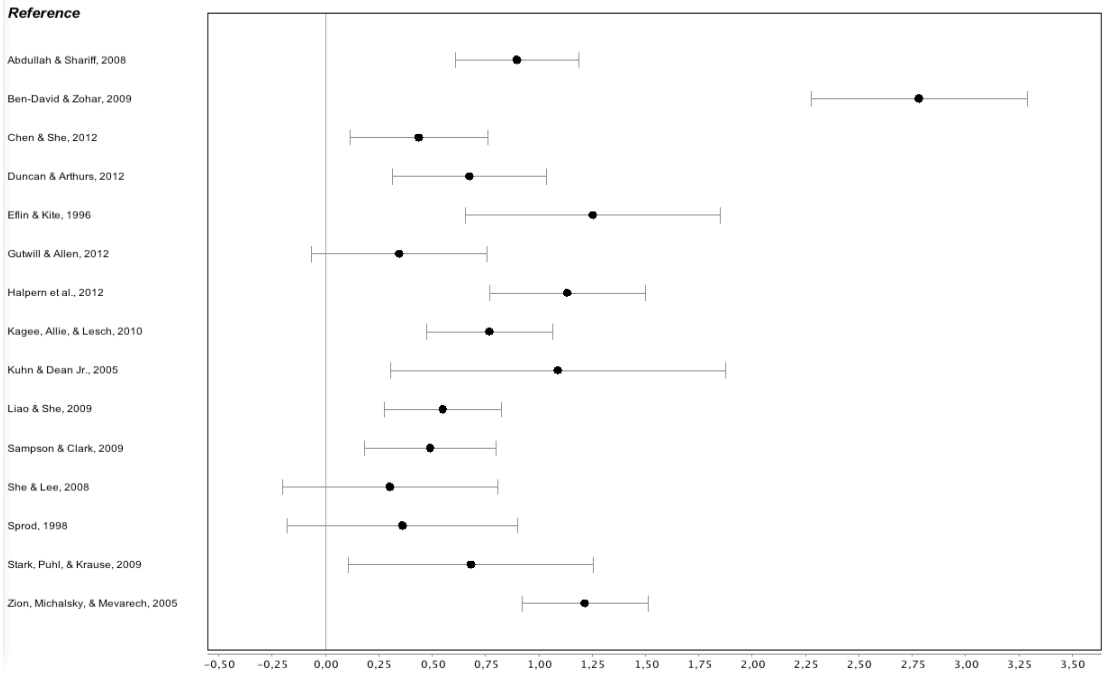


Figure 1. Forest plot of the effects included in the meta-analysis, indicating mean effect size and CI.

The second research question was: If there is variability of effect sizes across studies, to what extent do the content of scientific reasoning, the knowledge type, the technological support, and the type of learning activities included *in the intervention* explain the variability of the effects between studies?

The descriptive results of the subgroup comparison for the moderator variables in the intervention are shown in Table 1. The result of the subgroup comparison for the content of scientific reasoning revealed a highly significant difference between the groups ($Q(2) = 14.02$, $p < 0.01$). Interventions targeting the processes of scientific reasoning yielded larger effects than interventions targeting understanding science. Interventions on scientific argumentation yielded the lowest effects. The result of the subgroup comparison for the knowledge type also revealed a highly significant difference between the groups ($Q(2) = 32.01$, $p < 0.01$). The largest effects were found in interventions that included the combination of knowledge types; the lowest effect in interventions that included processes and strategies. The result of the subgroup comparison for the technological support revealed no significant difference between presence or absence of technological support ($Q(1) = 0.21$, $p > 0.05$). Furthermore, no significant difference was found between the groups after adding the differentiation between technology in the context of constructivist pedagogy and technology in the context of other pedagogies ($Q(2) = .37$, $p > 0.05$). The results of the subgroup comparison for the learning activities revealed a highly significant difference between constructive and interactive activities ($Q(1) = 17.68$, $p < 0.01$) and also highly significant differences after including the differentiation between instructional and peer dialog into the moderator analysis ($Q(2) = 18.77$, $p < 0.01$). interventions utilizing constructive activities yielded larger effects than interventions utilizing interactive activities. Moreover, the addition of the differentiation between peer dialog and instructional dialog increased the amount of explained variability between the studies.

Table 1: Descriptive results of the moderator analysis of the intervention

	Number of studies	g	Lower bound of the CI	Upper bound of the CI
Content of scientific reasoning				
Processes of scientific reasoning	3	1.28	0.99	1.58
Scientific argumentation	6	0.65	0.50	0.81
Understanding Science	3	0.84	0.65	1.04
Knowledge type				
Facts and concepts	5	0.73	0.56	0.90
Processes and strategies	5	0.46	0.29	0.64
Combination	4	1.20	1.01	1.38
No technological support	*6*	*0.83*	*0.67*	*0.99*
Technological support	*9*	*0.78*	*0.66*	*0.90*
Constructivist pedagogy	6	0.80	0.66	0.93
Other pedagogies	3	0.73	0.45	1.02

	Number of studies	g	Lower bound of the CI	Upper bound of the CI
Learning activity				
Active	0			
Constructive	5	1.09	0.93	1.25
Interactive	9	0.66	0.53	0.78
Interactive: instructional dialog	4	0.74	0.54	0.94
Interactive: peer dialog	5	0.60	0.44	0.76

The third research question was*:* If there is variability across studies, to what extent do the content of scientific reasoning, the knowledge type, and the transfer with respect to different knowledge types *in the post-test* explain the variability between studies?

The descriptive results of the subgroup comparison for the moderator variables in the post-test are shown in Table 2. The result of the subgroup comparison for the content of scientific reasoning revealed a highly significant difference between the groups ($Q(2) = 23.28$, $p < 0.01$). The inclusion of processes of scientific reasoning in post-tests yielded higher effects than understanding science. Post-test measuring scientific argumentation yielded the lowest effects. The result of the subgroup comparison for the knowledge type also revealed a highly significant difference between the groups ($Q(2) = 10.53$, $p < 0.01$). The largest effects were found in post-tests that measured facts and concepts, the lowest effect was found in post-tests that measured a combination of knowledge types. The result of the subgroup comparison for the transfer with respect to different knowledge types also revealed a highly significant difference between the groups ($Q(1) = 15.75$, $p < 0.01$). Post-tests that included no transfer yielded the largest effects than post-tests that included transfer.

Table 2: Descriptive results of the moderator analysis of the post-test

	Number of studies	g	Lower bound of the CI	Upper bound of the CI
Content of scientific reasoning				
Processes of scientific reasoning	4	1.25	1.04	1.46
Scientific argumentation	3	0.56	0.35	0.77
Understanding Science	6	0.75	0.61	0.88
Knowledge type				
Facts and concepts	5	1.03	0.85	1.20
Processes and strategies	4	0.78	0.59	0.96
Combination	4	0.61	0.43	0.79
Transfer with respect to different knowledge types				
No transfer	10	0.94	0.82	1.06
Transfer	3	0.47	0.28	0.67

Discussion

This meta-analysis carried out a systematic examination of interventions that aimed to foster scientific reasoning and found a large mean effect, suggesting that scientific reasoning can be successfully fostered by interventions. Furthermore, some variability between the studies may be explained by variables in the intervention and variables in the post-test of scientific reasoning.

Studies in which the interventions focused on processes of scientific reasoning such as hypothesis generating, experimenting, and evidence evaluation showed a larger mean effect than studies that focused on an understanding of science; studies aimed at facilitating scientific argumentation had the smallest effect size. The same pattern was found in the post-test of scientific reasoning. Measuring processes of scientific reasoning reached a larger mean effect than than measuring an understanding of science, while measuring scientific argumentation yielded the smallest effect size. The knowledge type also played an important role in intervention and post-test. The largest mean effect was achieved in studies that included a combination of knowledge types in the intervention. Regarding the post-test of scientific reasoning, the largest effects were achieved when only facts and concepts were measured.

The effects of the moderator analyses have to be cautiously interpreted, however, because the tests of significance refer to subgroup comparisons and do not provide a test of significance for comparisons between single variables. The moderator analysis shows, for instance, a significant difference between contents of scientific reasoning in interventions. However, it does not provide a test of significance between *processes of scientific reasoning* and *understanding science*.

The results of the moderator analysis with respect to technological support in the learning environment do not support the assumption that technology can increase the effect size, even when combined with a constructivist pedagogy (Rosen & Salomon, 2007).

The differentiation in learning activities between interventions revealed a significant difference, validating to an extent the distinction put forward in the ICAP hypothesis by Chi (2009). However, the assumption that *interactive* activities are more effective than *constructive* activities is not supported in the context of intervention studies on scientific reasoning. The larger effect of constructive, in comparison with interactive, activities might be explained by the topic of the intervention. It would be reasonable to assume that interactive activities, especially with peers, provide both aspects that support the learning process and aspects that are detrimental. The evidence provided by Chi (2009), supporting the hypothesis that interactive are more beneficial than constructive activities, compares situations only in which the interactive activity is similar to the constructive one, except for the presence of another individual. In this comparison, further aspects are neglected that occur in interactive activities at times; add to the *cost* side; and might negatively affect the learning process; such as time that is spent on coordination and eristic arguments. Scientific reasoning is a complex target for intervention; thus, the *cost* side of interactive activities might have become more influential. If this explanation is valid, ICAP needs to be differentiated with respect to potential collaboration costs that might be higher in more complex reasoning tasks without additional guidance. Future analyses could more closely examine the interaction process itself to test this modification of ICAP. In addition, future analyses should try to include the *active* category of activities in learning environments for scientific reasoning and argumentation, to comprehensively test the validity of ICAP in this context.

The results of the moderator analysis concerning the transfer suggest that the range of transfer between intervention and post-test explains some variability between the studies. The absence of transfer yielded larger effects than the presence transfer. This result is coherent with conceptualizations of transfer (e.g. Barnett and Ceci, 2002) and could additionally be interpreted as a validation of the sample of studies that was included in the meta-analysis.

Even though most results of the meta-analysis are highly significant, interpretations of the results have to be made cautiously. The main limitations of the meta-analysis directly result from the selection of studies. As Eisend (2004) points out, the research and publication process favors studies that report significant results, leading to a bias in the published studies. Furthermore, the sample included in the meta-analysis was limited to studies that included scientific reasoning and related search terms in the title, which might have enhanced a bias in favour of studies that included an successful intervention for scientific reasoning.

This meta-analysis provides a first overview of intervention studies concerning scientific reasoning and shows a large effect of intervention on scientific reasoning. Furthermore, we were able to identify moderator variables in the content, pedagogy, and assessment. Interventions for scientific reasoning which engage learners in constructive activities are more successful than interventions which engage learners in interactive activities; this result validates the distinction made by Chi (2009) but, at the same time, disconfirms the order of the activities. Furthermore, the type of content and knowledge fostered in the intervention, and measured in the post-test, influences the success of the intervention. Further research is needed to test directly the effects found in this meta-analysis.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

*Abdullah, S., & Shariff, A. (2008). The effects of inquiry-based computer simulation with cooperative learning on scientific thinking and conceptual understanding of gas laws. *EURASIA Journal of Mathematics, Science & Technology Education, 4*(4), 387-398.

American Association for the Advancement of Science. (1989). *Project 2061- Science for all Americans*. Washington- AAAS.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.

*Ben-David, A., & Zohar, A. (2009). Contribution of meta-strategic knowledge to scientific inquiry learning. *International Journal of Science Education, 31*(12), 1657-1682. .

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research*, *80*(3), 336-371.

*Chen, C.-H., & She, H.-C. (2012). The impact of recurrent on-line synchronous scientific argumentation on students' argumentation and conceptual change. *Educational Technology & Society, 15*(1), 197-210.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105.

diSessa, A. A. (2008). A "theory bite" on the meaning of scientific inquiry: A companion to Kuhn and Pease. *Cognition and Instruction*, *26*(4), 560-566.

*Duncan, D. K., & Arthurs, L. (2012). Improving student attitudes about learning science and student scientific reasoning skills. *Astronomy Education Review,* 11(1), 010102–1–010102–11.
*Eflin, J. T., & Kite, M. E. (1996). Teaching scientific reasoning through attribution theory. *Teaching of Psychology, 23*(2), 87.
Eisend, M. (2004). *Meta-analyse – Einführung und kritische diskussion*, Discussion Paper of the Economics Department of the Free University Berlin, No. 2004/8.
Giere, R. N. (1979). *Understanding Scientific Reasoning*. New York,NY: Holt, Rinehart & Winston.
*Gutwill, J. P., & Allen, S. (2012). Deepening students' scientific inquiry skills during a science museum field trip. *Journal of the Learning Sciences*, *21*(1), 130-181.
*Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity, 7*(2), 93-100.
*Kagee, A., Allie, S., & Lesch, A. (2010). Effect of a course in research methods on scientific thinking among psychology students. *South African Journal of Psychology, 40*(3), 272-281.
Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive science*, 48, 1-48.
Kuhn, D. (1991). *The skill of argument.* New York,NY: Cambridge University Press.
Kuhn, D. (1993). Science as argument: implications for teaching and learning scientific thinking. *Science Education, 77*(3), 310-337.
Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, *94*(5), 810-824..
*Kuhn, D., & Dean Jr., D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science (Wiley-Blackwell), 16*(11), 866-870.
Lawson, A. E. (1995). *Science teaching and the development of thinking.* Belmont, CA: Wadsworth.
*Liao, Y.-W., & She, H.-C. (2009). Enhancing eight grade students' scientific conceptual change and scientific reasoning through a web-based learning program. *Educational Technology & Society, 12*(4), 228-240.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks,CA: Sage.
Mayer, R. E. (2012). Information processing. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 85-99). Washington, DC: American Psychological Association.
Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, *328*(April), 463-467.
Rosen, Y., & Salomon, G. (2007). The differential learning achievements of constructivist technology -intensive learning environments as compared with traditional ones: A meta-analysis. *Journal of Educational Computing Research, 36*(1), 1-14.
*Sampson, V., & Clark, D. (2009). The impact of collaboration on the outcomes of scientific argumentation. *Science Education, 93*(3), 448-484.
*She, H.-C., & Lee, C. Q. (2008). SCCR digital learning system for scientific conceptual change and scientific reasoning. *Computers & Education, 51*(2), 724-742.
*Sprod, T. (1998). "I can change your opinion on that": Social constructivist whole class discussions and their effect on scientific reasoning. *Research in Science Education, 28*(4), 463-480.
*Stark, R., Puhl, T., & Krause, U.-M. (2009). Improving scientific argumentation skills by a problem-based learning environment: effects of an elaboration tool and relevance of student characteristics. *Evaluation & Research in Education, 22*(1), 51-68.
Wilson, D. B. (2005). *Meta-analysis macros for SAS, SPSS, and Stata*. Retrieved, June, 16, 2013, from http://mason.gmu.edu/~dwilsonb/ma.html
Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*(1), 99-149.
Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27*(2), 172-223.
*Zion, M., Michalsky, T., & Mevarech, Z. R. (2005). The effects of meta-cognitive instruction embedded within an asynchronous learning network on scientific inquiry skills. Research Report. *International Journal of Science Education, 27*(8), 957-983.

Knowledge Organization with Multiple External Representations for Socioscientific Argumentation: A Case on Nuclear Energy

Bahadir Namdar, The University of Georgia, Athens, GA, USA, baha@uga.edu
Ji Shen, University of Miami, Coral Gables, FL, USA, j.shen@miami.edu

Abstract: Given the vast amount of information readily available through the Internet in different forms of representation, learners need to organize their knowledge strategically in order to make a sound argument, especially on a complex topic such as socioscientific issues. In this study, we designed a science unit on nuclear energy using a newly developed online knowledge and learning management system that offers three types of representations: pictorial, textual, and concept maps. We investigated how learners organized their knowledge with multiple external representations and how their knowledge organization practices interacted with their argumentation. Our results indicated that concept maps and wiki entries were more connected than the pictorial modes in the network of knowledge entries created by the learners. Moreover, we found evidence showing that students' knowledge organization enabled them to draw information while arguing and that their argumentation guided them to advance their knowledge organization practices.

Introduction

Socioscientific issues (SSI) are useful pedagogical tools in science classrooms. They are often relevant to learners' lives and therefore, can promote interest in learning science (Kolstø, 2006; Sadler & Zeidler, 2005). They can engage diverse learners in participating in scientific discourse and argumentation (Zeidler & Nichols, 2009). Producing sound arguments on SSI requires learners to synthesize complex information and use evidence-based reasoning (Zeidler & Nichols, 2009). It is important to create opportunities for learners to discover mutual influence of science and society, which can empower learners' science learning as well as their decision making in their lives(Simonneaux, 2008). Zeidler & Nichols (2009) submitted that arguing and debating about SSI in classroom could engage students in scientific thinking and reasoning, and provide students with firsthand experience of the advancement of scientific knowledge in daily life.

With the rapid development of information and communication technology (ICT), relevant (and irrelevant) information and data pertinent to particular SSI are distributed across a vast network of resources. Such information and data are often presented in multiple external representations (MER), such as tables, graphs, texts, models, and pictures (Ainsworth, 2006; Chandrasegaran, Treagust, & Mocerino, 2011). When arguing about a given SSI, students need to organize relevant information in an effective way and construct and represent their arguments accordingly, especially in computer supported collaborative learning (CSCL) environments. Therefore, learners must know how to search, sort, cluster, tag information in the forms of representations that reflect their understanding. We call this process *knowledge organization* (Namdar & Shen, 2013).

Although SSI, MER, and argumentation have been widely studied in science education, the process of knowledge organization with MER and the interaction between knowledge organization and argumentation on SSI remains relatively unexplored. Hence, to address this gap in the literature, we designed a learning unit on a SSI and asked learners to organize their knowledge using MER and argue about the issue using a CSCL platform. Our inquiry has focused on the following two research questions: (In an argumentation-based CSCL environment)

1. How do learners organize knowledge effectively with MERs?
2. How does learners' knowledge organization with MER interact with their argumentation practices?

Theoretical Framework

This study is built on both cognitive and social constructivist perspectives toward learning science. First, from a social constructivism standpoint, we hold the assumption that learning occurs within a community of learners and that learners construct their understanding within this socially interactive context (Lemke, 2001; Vygotskiĭ , 1986). Therefore, in our research we created a collaborative learning environment for learners to construct their knowledge using a CSCL tool in small groups and also at the class level. Second, based on the assumption of cognitive constructivism, learners use tools to construct their individual understanding (Piaget, 1970). In this study students used different representations to coordinate and construct meanings of a SSI.

Researchers have reported many benefits of using MERs in science instruction, such as capturing learners' interest (Ainsworth, 1999) and enhancing students' understanding of scientific concepts (Chandrasegaran et al., 2011; Waldrip, Prain, & Carolan, 2010). According to Ainsworth (1999), one goal of utilizing MER is to enable learners to take advantage of the benefits that are associated with each representation,

 © ISLS

and help individuals with varying learning abilities. It has been noted that naturally many students have meta-representational competency (diSessa, 2004), which might lead to improved learning gains (Grossen & Carnine, 1990).

Ainsworth (2004, 2006) argued that learners have difficulties with integrating information from multiple data sources as they see individual representations in MERs in isolation. Therefore, learners need to make meaningful links between those representations. We define knowledge organization as "the process of searching, sorting, clustering, archiving, and externalizing knowledge in a systematic way to achieve a better understanding of the world" (Namdar & Shen, 2013, p. 345). This process will enable learners to have more systematically organized and externalized information which will help them to retrieve information from multiple sources.

Our conceptualization of knowledge organization also stems from knowledge integration and knowledge building theories. We adopt the premises of helping learners make connections about their ideas and conceptually integrate those ideas from knowledge integration (Linn, 2006). We also believe in the needs of creating a knowledge building community and representing knowledge as in the forms of epistemic artifacts (Sterelny, 2005) from the knowledge building perspective (Scardamalia & Bereiter, 2006). Knowledge organization therefore embraces two distinct attributes. First, knowledge organization requires learners to actively create, manipulate, and connect MERs, especially with the aid of technological tools. Second, knowledge organization demands the creation of interrelated knowledge webs at different social levels.

Learning Platform and Unit Design

Innovative Knowledge Organization System (iKOS)

iKOS is a web-based knowledge organization platform that helps learners to organize knowledge both individually and collectively (ikos.miami.edu). It follows three core design principles: providing learners with means to externalize knowledge in multiple forms and facilitate the transformation among them; engaging learners' knowledge organization and construction at both the private and the public spaces and ease the flow between them; nurture learners' independent, critical thinking as well as collaborative mind set and catalyze the transition between them. In iKOS, learners can create knowledge entries in three distinct representation modes: *Event, Wiki,* and *Concept Map*. In Event, learners search the web, find, and upload pictures of a complex scientific phenomenon. Learners tag and annotate those pictures to understand the phenomenon of interest. Wiki is similar to the Wikipedia interface in which learners can write text. Learners can also create Concept Maps in the system and visualize the connections among a set of related science concepts (Novak & Cañas, 2008). iKOS automatically links and visualizes learners' knowledge entries based on similar keywords.

Context and Lesson Sequence

This study took place in a large southeastern public US university. The sample included a class of 23 pre-service teachers who were taking a middle school methods class for science teaching. The study was conducted in 4 sections in total of six hours. The first, second, and last sections took one hour, while the third section took three hours. We assigned students in four groups and each group had five students except one group with four students.

Although we focus on the learning aspect in the study, there are two main reasons to recruit pre-service teachers. First, we concur with Zeidler, et al.(2002) in that "[pre-service teachers] are in a position for effecting change with the future learners they teach concerning the topics that have been identified as seminal issues for science education" (p.346). Therefore, introducing an innovative learning approach to pre-service teachers may be more transformative for the future of science education. Second, introducing a new learning approach, especially in its early phase, may be risky for existing teachers as they face the pressure of high-stakes testing. Pre-service teachers are more accessible in this regard.

The lesson sequence was implemented as the following: (1) introduction to argumentation and concept mapping; (2) Introduction to the topic of nuclear energy by reading an article and watching videos that focused on pros and cons of nuclear energy; (3) Creating iKOS entries individually; (4) Creating iKOS entries collaboratively in small groups on a particular scientific aspect of nuclear energy and engaging in argumentation; (5) Peer critique and revision of MERs, 6) Final presentation and argumentation. At the end of the unit, students presented their findings and argued for their stance on the issue of building nuclear power plants. First author taught the class sessions for this study.

Data Collection and Analysis

We employed a *concurrent nested* mixed methods study design (Creswell, Plano Clark, Gutmann, & Hanson, 2003) to *better understand* the knowledge organization and argumentation practices (Greene, 2007). We collected multiple sources of data for the purposes of complementarity as the different data were used to tap into

different facets of the learning process (Greene, Caracelli, & Graham, 1989). The data collection included participant observation of class interventions (Suzuki, Ahluwalia, Arora, & Mattis, 2007), learners' artifacts (MERs created on iKOS), log file created by iKOS, and video recordings of classes. Quantitative methods that focused on understanding the students' knowledge organization through the use of MERs and qualitative methods aimed at understanding the interaction between knowledge organization and argumentation practices (Greene et al., 1989; Greene, 2007). Both qualitative and quantitative data were collected concurrently.

The first research question was answered by analyzing the log file that was generated by iKOS. As the iKOS system automatically interlinks MERs that were created by the students, this file reports the number of links, the number of entries created, and the number of team members if the entry is co-created. We considered each individual iKOS entry as a node (or actor) in a graph to enact social network analysis (Knoke & Yang, 2007). We calculated the mean normalized degree centrality by adding all the links that were associated with one entry and divided this number by the possible number of links that this entry could possess in connection to the knowledge web (Knoke & Yang, 2007). Then, we added all the individual mean normalized degree centralities that were associated with one mode of MER and normalized this number by dividing it by the possible number of links that entries of particular mode have in the network.

To have a better understanding of the actors in the network, we ran key actor analysis using the *R* statistical package. We calculated betweenness centrality and eigenvector centrality measures for each actor in the network. Betweenness centrality measured the number of shortest paths that an actor is on, which makes this actor important in controlling the flow of the information in the network (Knoke & Yang, 2007). Eigenvector centrality measured how central an actor is and how central the ties of this actor are in the network (Bonacich, 2007). These two measures indicate how well connected an actor is in the network.

To answer our second research question, based on the results of key actor analysis, we examined the argumentation that involved those students who created the highest ranked key actors. We analyzed the videos by adopting the analytical model suggested by Powell et al. (2003). First, we watched the videotapes several times to become familiar with the content of the video without intentionally imposing critical and theoretical lenses toward the data. Second, based on 5-minute intervals, we described data without including interpretations or inferences. After becoming familiar with the content of the videos, we carefully identified the significant instances, or critical events (Powell et al. 2003). Critical events in this research refer to the instances in which learners argue collaboratively and take actions in their knowledge organization (revising, editing entries, creating new ones) or use their knowledge organization entries in their arguments. Next, we transcribed relevant argumentation sessions. In the coding stage, we specifically examined students' statements and their relationship to the specific entries that the students created or how the students' arguments led them to create a certain entry type. For instance, when a student wrote in her wiki "The fact that this power plant will be located 26 miles from SkyCity makes it a potential hazard to a large population of people" (Haley, Wiki entry) and employed a similar idea in her argumentation in the class, we coded her verbal argument as "a reference to a wiki."

Results

Network of Students' iKOS Entries

The students created 17 events, 23 wikis, and 20 concept map entries (see Table1). Our results indicated that the wiki mode had the highest mean degree group centrality. This result suggests that in this sample the most centralized/prominent entry mode was the wiki mode. These results also show that the event mode was the least centralized. Moreover, normalized degree centrality for the concept map mode was close to that of the wiki mode.

Table 1. Mean Normalized Degree Centralities for iKOS Modes

Entry Mode	Total Number of Entries	Normalized Degree Centrality	*SD*
Event	17	0.38	0.28
Wiki	23	0.51	0.21
Concept Map	20	0.49	0.22

Figure 1 shows the results of our key actor analysis weighted by eigenvector and betweenness centralities. Each numbered actor in the figure represents an iKOS entry. The diameter of each actor is proportional to the product of eigenvector and betweenness centralities. We closely examined the top five key actors in the network (#A149, A145, A130, A121, A157; see figure 1). Our results indicated that three out of five actors were concept maps and the remaining two actors were wiki entries.

Additionally, the network density was 0.55which means that 55% of the possible links were present. In other words, this knowledge network was a very dense network.

Interpretation

Overall, the knowledge web that was generated as a result of this intervention was highly connected (see Figure 1), which suggests a good knowledge organization practice by the students. The high density of the network also suggests that the students were mindful when organizing their knowledge; that is, they carefully connected concepts in their concept map entries and generated keywords for their wiki entries as the tags and keywords are the sources for links in iKOS.

Both concept maps and wiki entries were more central than the event entries, as expected. Since the function of a concept map is to connect big ideas (Novak & Cañas, 2008), it makes sense for them to be central, relatively speaking, in a knowledge web. Also, students created more wiki entries. This was natural as textual representations are more approximate to verbal arguments. This result is consistent with our SSI unit that had an argumentation-based design. On the other hand, the event entries were not as connected in the knowledge web (Table 1). An event entry was relatively harder to create compared to the wiki and concept map entries in the system. Also, the function of an event entry is to encourage students to reflect their understanding of scientific processes and concepts.

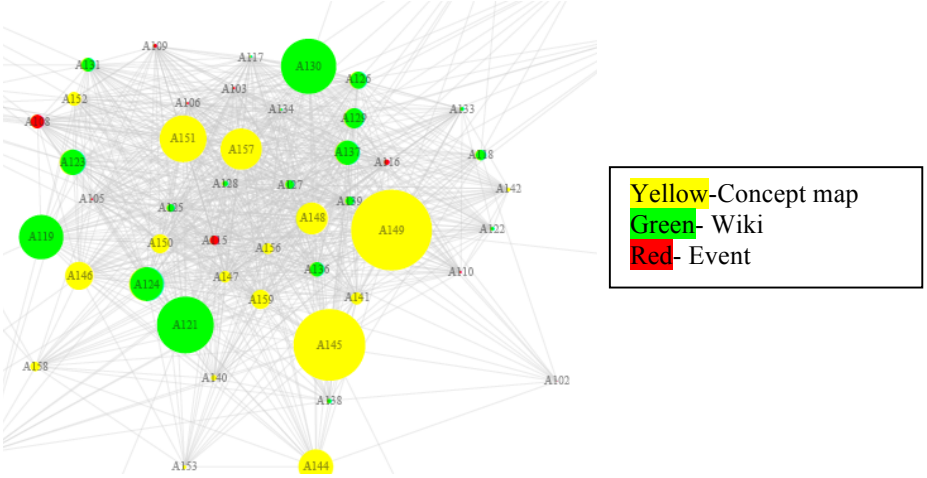


Figure 1. Key Actor Analysis for iKOS Entries Weighted by Betweenness and Eigenvector Centralities

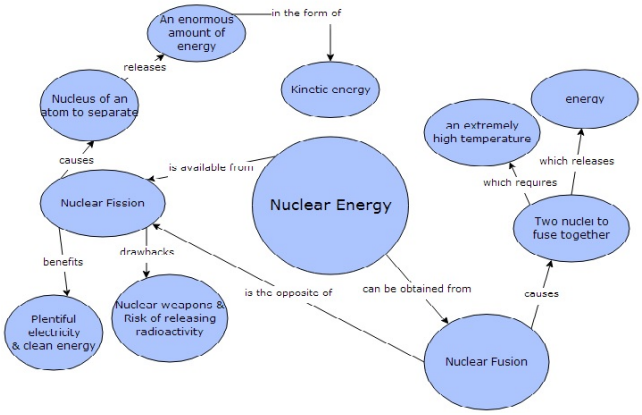


Figure 2. Haley's Concept Map Entry

Interaction between Knowledge Organization and Argumentation: The Case of Haley

To address our research question 2, we further examined the top ranked key actors. The concept map entry created by Katie was the top ranked entry overall. However, Katie needed to leave the class after the first group discussion. Therefore, we decided to focus on the case of Haley, who created the second highest ranked key actor, which was a concept map (see Figure 2).

In her homework Haley generated an entry in each mode (step (3) in lesson sequence). In her event entry, Haley inserted a picture of a nuclear power plant and focused on how this nuclear fission power plant worked. In her concept map, she looked at different types of nuclear energy creation processes: nuclear fission and fusion (See Figure 2). In her long wiki entry she listed the benefits, drawbacks of nuclear energy, and wrote her claim to the question of 'should we build nuclear power plant in our state?" She summarized her *claim* as:

"no, we should not build nuclear power plants in our state." and listed six pieces of *evidence* : "(1) Costs billions of dollars, (2) Benefits do not outweigh the risks, (3) Power companies don't pay if accident happens, families pay by losing their loved ones (sometimes years later due to cancer), (4) Exposure to radiation kills living cells, (5) People around power plant are displaced (property value goes down since no one wants to live near a power plant), (6) Located 26 miles northeast of a city - too close to a densely populated area." Finally, she wrote her justification:

> While nuclear energy is a relatively clean source of energy, the potential release of radiation from nuclear power plants is a huge risk. Accidents can happen, and they have happened. The fact that this power plant will be located 26 miles from SkyCity makes it a potential hazard to a large population of people. We do not have the technology yet to make sure harmful radiation is not released. Why should we build a nuclear power plant when the proper safety measures are not in place? [Haley, wiki entry, 04/17/2013]

We compared Haley's individual entries with those created by her teammates' (her group had 5 students). Two students did not create any entries prior to the class. None of the other two students created a concept map. Haley was the only student in her group who created all three types of external representations. When we look at the event entries, we saw that all three students (Haley, Rachel, and Melissa) inserted a picture of a nuclear power plant. Melissa only inserted a picture and did not use any tags, a poor knowledge organization practice. Rachel on the other hand, used seven tags in her event. But instead of using central scientific concepts such as radioactive elements, heat, and energy as Haley did, she used specific terms such as coolant and cooled liquid to refer to concrete objects.

Key actor analysis indicated that Haley's event entry was more central compared to the other two students' event entries. For wiki entries, Melissa's wiki entry was one of the key actors (A130). In her entry she wrote:

> Nuclear energy comes from a process called fission. This generates heat which produces steam, which then generates electricity. There is another type of reaction called fusion (Melissa, Wiki Entry, 04/17/2013).

Rachel, on the other hand, wrote several things about nuclear energy (e.g., the percentage of energy and electricity generated from nuclear power, some benefits such as reducing carbon emissions and some drawbacks such as radiation threats to environment).

When coming to the class, students were asked to argue on the following two questions in their small groups (step (4) in lesson sequence): Should we build nuclear power plants in our state? For how long should we depend on nuclear energy as an alternative energy source? Haley immediately initiated the discussion:

> [Excerpt 1. Transcripts 09/01/2013: 04:35-05:32]
>
> Haley: I do not think we should, because in the article it is talking about building it like within twenty-five miles, or something like that from SkyCity. So if anything were to go wrong. That would affect… it's like very close to a high population of people.
>
> Ashley: Yeah.
>
> Haley: That would affect a lot of people.
>
> Researcher: What would be the effect?
>
> Haley: Well, the radiation kills living cells so like that might not affect them right away but it could affect them like over ten years if they are exposed to that radiation, and it develops cancer and then ten years down the line, power plants are not gonna be the ones that wait for that. They are not gonna come and say "oh you know we gave you cancer" and we are not gonna be able to fix it so people are gonna lose their lives.

This excerpt indicates that Haley was able to draw information from the wiki entry she created in which she mentioned the "lethal amounts of radiation, "the potential release of radiation from nuclear power plants is a huge risk," and "The fact that this power plant will be located 26 miles from SkyCity makes it a potential hazard to a large population of people."

As Haley initiated the discussion, she actually dominated it. She drew information seven times from her wiki, four times from her concept map, and one time from her event entry. Following their group argumentation session Haley gave an example of a movie that she saw in which a lawyer was trying to save a community that lived near a nuclear power plant. It contaminated the ground water and caused cancer in the

community. After giving this example Haley drew the group's conversation to a certain point that they started to talk about cancer and radiation.

[Excerpt 2. 09/01/2013: 15:30-15: 41]
Haley: It [nuclear power plant] contaminates…even if there isn't an accident they are still spreading nuclear radiation into the ground.
Rachel: They cause cancer.
Haley: Yeah, it caused cancer and they did not know where it was coming from and they just thought, you know, cancer, you really never know where it comes from.

After about ten minutes into their small group discussion, students were asked to choose one scientific aspect associated with nuclear energy and create iKOS entries to explain it. Haley's group decided to focus on the topic of radiation exposure and its connection to cancer (see excerpt 3 below).

[Excerpt 3. 09/01/2013: 17:35- 18:05]
Ashley: Why do not we do radiation and cancer, so we do not want it.
Rachel: [Inaudible]
Ashley: We are done. We are gonna do radiation and cancer
Haley: How radiation kills cells?
Ashley: Yeah.
Haley: Radiation exposure and cancer.

As a result, Haley's group created one entry on each mode and all entries focused on the connection between radiation and cancer. However, there was a slight difference in terms of the information represented in three modes. In the wiki, they stated that **"**The radiation that comes from nuclear reactors is ionizing radiation," and listed two dangers:

(1) It can damage DNA leading to mutations, thus potentially causing cancer or death of the cell. Damage to the cell can take place in less than a second, but cancer can take years to develop, and
(2) Ionizing radiation can be more carcinogenic than other types of radiation, and lead to cancers such as: thyroid, bone marrow, leukemia, skin, lung, stomach, breast, etc.

They also wrote about the dangers of exposure to radiation and the testing of nuclear reactors in their wiki. In the event entry, they inserted several pictures that showed different information: about how UV photon mutates the DNA, how normal cells mutate to cancer cells, and how cancer cells leads to a tumor. In the concept map, they summarized the types of radiation and tied those to the cause of cancer.

Interpretation

Our analysis showed that Haley was able to draw information from different representations as she argued on the nuclear energy topic within her group. She mostly used text-based entries (wiki-7 times; concept maps-4 times) as the source of information for her argument. This is natural as she was asked to verbally argue on the same topic, and she referred to her statements in the wiki mode. She only referred once to her event entry. Acquiring information from a pictorial representation during a verbal argumentation might be more challenging, as it requires students to explain and elaborate on the nuances that the picture includes. Through looking into Haley's case, we extracted two characteristics of the type of knowledge organization practices that contribute to argumentation.

The first characteristic of a better knowledge organization might be the number of different types of representations that were created. Compared to her group members Haley was the only student who created all three representations. The creation of more representations might have allowed her to acquire different pieces of information as she used MERs to represent slightly different units of information on a given topic.

The second characteristic of better knowledge organization that fostered argumentation speaks to the quality of the entries that were created. For instance, although Melissa's wiki was a key actor in the knowledge network, we see that she did not elaborate on the fission concept. The reason for Melissa's wiki entry to be key actor is that it connects with key actors in the knowledge network (e.g., #A148, A149, A157, A159; see figure 1); as Melissa used the key concepts as her keywords (i.e., nuclear, energy, nuclear energy, and fission). Therefore, her entry was highly connected with other entries in the class. However, in the argumentation, she was not actively involved, and she did not acquire any information from her event entry during argumentation. Rachel, on the other hand, only listed facts about nuclear energy in her wiki entry without further elaboration. However, the structure of Haley's wiki entry was very organized. She wrote her claim and justification about

nuclear energy. As we also analyzed the group discussion, we saw that Haley was drawing information easily from her wiki entries for her arguments. Therefore, good knowledge organization might have fostered argumentation in the classroom. Similarly, as knowledge organization fostered students' argumentation, their argumentation also encouraged them to find and organize new information before they presented their findings to the class. However, specific types of interactions need further research Individually Haley's representations focused on different aspects of nuclear energy, but in their group knowledge organization, they created representations in three modes that focused on the connection between radiation and cancer. Additionally, as Haley used key concepts such as fission, fusion, and energy as central concepts in her concept map, her concept map was highly connected to other entries.

Ainsworth (2006) listed three functions of MERs: complementary roles, constrain interpretation, and construction of deeper understanding. As we see from Haley's individual entries, they were clearly complementary to each other. In that regard, MERs served to expand the scope of her understanding. Her group's entries, on the contrary, included more or less similar information. This was partially because in our lesson design we asked students to focus on one topic and present their findings to the class so that they might chose to create similar entries. On the other hand, Haley's group created their entries on the similar topic to emphasize their shared points. Therefore, to a certain degree, these similar entries in different modes served as constrains to each other, and helped the group develop a more in-depth understanding.

Conclusion and Discussion

Linn (2000) argued that "the internet provides a rich, confusing, chaotic, informative, persuasive set of scientific information" (p. 785). This is more so for SSI topics. Therefore, it is crucial to help learners use tools to organize and coordinate pieces of information that they use to make sound arguments. We designed a science unit on nuclear energy using an online knowledge and learning management system (iKOS) that offers three types of representations: pictorial, textual, and concept maps. Our results suggest that the interaction between knowledge organization and argumentation is bidirectional (i.e., students' knowledge organization enables them to draw information while arguing and that their argumentation guides them to advance their knowledge organization practices. On the other hand, we need more sophisticated methods to be able to capture relevant learning process both on an individual and a collaborative level.

According to Erduran and Evagorou (2012), educators should make the best use of visual representations as these are fundamental objects of enhancement of scientific knowledge and students' lives. Along these lines, we found that use of visual representations, MER in our case, enabled learners to organize their knowledge on SSI and helped them in argumentation. More specifically, our results indicated that when organizing knowledge with MERs, learners mostly created wiki and concept map entries. Additionally, when our system automatically created a knowledge-web based on the MERs that were created by the students, wiki and concept map entries were highly connected representation types. On the other hand, when we looked at the key representation across the whole network and the student (Haley) who created it, we also found that Haley was acquiring information from her wiki and concept map entries while she was arguing on the topic. Also, students' argumentation encouraged them to (re)organize knowledge.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2-3), 131–152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. doi:10.1016/j.learninstruc.2006.03.001

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*(4), 555–564. doi:10.1016/j.socnet.2007.04.002

Chandrasegaran, B. A. L., Treagust, D. F., & Mocerino, M. (2011). Facilitating high school students ' use of multiple representations to describe and explain simple chemical reactions. *Teaching Science*, *57*(4), 13–21.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. . (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, CA: Sage.

Erduran, S., & Evagorou, M. (2012). Visualising evidence and scientific methods : Implications for science education. In *National Association of Research in Science Teaching, Indianapolis, IN* (pp. 1–15).

Greene, J. C. (2007). *Mixed methods in social inquiry* (p. 216). San Francisco: John Wiley & Sons.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274.

Grossen, B., & Carnine, D. (1990). Diagramming a logic strategy e Effects on difficult problem types and transfer. *Learning Disability Quarterly*, *13*(3), 168–182.

Knoke, D., & Yang, S. (2007). *Social network analysis (Quantitative applications in the social sciences)*. (T. Liao, Ed.) (p. 133). Sage Publications, Inc.

Kolstø, S. D. (2006). Patterns in students' argumentation confronted with a risk focused socio-scientific issue. *International Journal of Science Education*, *28*(14), 1689–1716. doi:10.1080/09500690600560878

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching*, *38*(3), 296–316. doi:10.1002/1098-2736(200103)38:3<296::AID-TEA1007>3.3.CO;2-I

Linn, M. C. (2000). Designing the knowledge integration environment. *International Journal of Science Education*, *22*(8), 781–796.

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. Sawyer, Keith (Ed.), *The Cambridge handbook of the learning sciences* (pp. 243–264). New York, NY: Cambridge University Press.

Linn, M. C., & Hsi, S. (2000). *Computers,teachers, peers: Science learning partners*. Mahwah,NJ: Lawrence Erlbaum Associates.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. (IHMC CmapTools Report). Retrived from: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptMaps.htm

Namdar, B. & Shen, J. (2013). Knowledge organization with multiple external representations in an argumentation based computer supported collaborative learning environment. In Rummel, N., Kapur, M., Nathan, M., & Puntambekar, S. (Eds.). *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 1 — Full Papers & Symposia*. International Society of the Learning Sciences. *pp*, 344-351.

Piaget, J. (1970). *Genetic epistemology*. New York: Columbia University Press.

Powell, A. B., Francisco, J. M., & Maher, C. a. (2003). An analytical model for studying the development of learners' mathematical ideas and reasoning using videotape data. *The Journal of Mathematical Behavior*, *22*(4), 405–435. doi:10.1016/j.jmathb.2003.09.002

Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching*, *42*(1), 112–138. doi:10.1002/tea.20042

Scardamalia, M., & Bereiter, C. (2006). Knowledge building : Theory , pedagogy , and technology. In K. Sawyer, R. (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97–115). New York, NY.

Simonneaux, L. (2008). Argumentation in socio-scientific contexts. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: An overview* (pp. 179–199). Springer Netherlands.

Sterelny, K. (2005). Externalism, epistemic artefacts and the extended mind. In R. Schantz (Ed.), *The Externalist Challenge: New Studies on Cognition and Intentionality*. Berlin: de Gruyter.

Suzuki, L. A., Ahluwalia, M. K., Arora, A. K., & Mattis, J. S. (2007). The pond you fish in setermines the fish you catch: Exploring strategies for qualitative data collection. *The Counseling Psychologist*, *35*(2), 295–327.

Vygotskiĭ , L. (1986). *Thought and language*. (A. Kozulin, Ed.). Ipswich, MA: MIT Press.

Waldrip, B., Prain, V., & Carolan, J. (2010). Using Multi-Modal Representations to Improve Learning in Junior Secondary Science. *Research in Science Education*, *40*(1), 65–80. doi:10.1007/s11165-009-9157-6

Zeidler, D. L., & Nichols, B. H. (2009). Socioscientific issues: Theory and practice. *Journal of Elementary Science Education*, *21*(2), 49–58. doi:10.1007/BF03173684

Zeidler, D. L., Walker, K. A., Ackett, W. A., & Simmons, M. L. (2002). Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education*, *86*(3), 343–367.

Students' Resources for the Construction of Scales for Graphing

Cesar Delgado, University of Texas at Austin, 1912 Speedway Stop D5500, Austin, TX 78712, USA, cesar_delgado@austin.utexas.edu
Margaret Lucero, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, USA, mlucero@scu.edu

Abstract: Graphing is fundamental in the scientific process. Scales are key but little-studied components of graphs. Using a "fine-grained constructivist" perspective, we investigated the resources undergraduate students activate in constructing a scale for difficult data sets ranging over 10 or more orders of magnitude. Following a constant comparison method, we identified resources including: show all elements, zoom in, ratio, halving, and conversion. Students also used their knowledge of bar graphs, histograms, and logarithmic, powers of ten, and linear scales as resources, at times inappropriately. Implications for instruction are outlined.

Student Resources for Scale Construction

Graphing is a fundamental part of the scientific process. Scales are key but little-researched components of graphs. We investigate the resources that undergraduate students activate in constructing a scale for difficult data sets, adopting a "fine-grained constructivist" perspective (Elby, 2000). We use difficult tasks to preclude students' simply following well-known procedures. Identifying students' resources for representation and building upon them in the classroom should lead to deeper understanding of conventional graphing practices.

Graphs and Scales

Natural phenomena can often be modeled mathematically; graphs can portray experimental data or the modeled relationship between variables. Graphs effectively describe continuous variation (Lemke, 1998) and succinctly summarize large amounts of data (Latour, 1987). US science standards note a role for graphing in scientific and engineering practices such as "Analyzing and interpreting data" (National Research Council, 2012).

Scales are essential components of graphs, and play an important role in graph interpretation (Leinhardt, Zaslavsky, & Stein, 1990). Scales permit the graphical representation of the magnitude of physical or mathematical quantities, or of the different values of a qualitative variable. Common scales consist of evenly spaced intervals portraying constant additive differences (linear scale), multiplicative differences (logarithmic or powers of ten scale), different values of a nominal- or ordinal-level variable (qualitative scale), or "bins" with ranges of data (histogram scale). In linear and logarithmic scales, the magnitude of data points is represented solely by their location along the scale; such scales are "homogeneous", and they allow the relative magnitude of data points to be directly compared from the graph (Nemirovsky & Tierney, 2001). Homogenous scales follow the convention that equal lengths along the scale represent an equal number of units (Leinhardt et al., 1990). The powers of ten scale is to be taught in high school, and linear, qualitative, and histogram scales in sixth grade or earlier, according to the US Common Core State Standards (NGACBP & CCSSO, 2010).

Research shows that students have difficulty in producing or interpreting conventional graphs (see review by Leinhardt et al., 1990). Difficulties constructing scales include placing quantitative data points on successive, evenly spaced tick marks regardless of values (e.g., Brasell, 1990), or taking a scale as discrete, i.e., considering that there are no points between the labeled points (Leinhardt et al., 1990).

Theoretical Framework: Fine-Grained Constructivism

Recent research has explored the resources students possess that enable them to understand and engage in representation (e.g., diSessa, Hammer, Sherin, & Kolpakowski, 1991; Elby, 2000; Hammer, Elby, Scherr, & Redish, 2005; Sherin, 2000). This research stresses the importance of building on and reorganizing prior knowledge. It conceives of learners as holding multiple, fine-grained understandings to be called into service as needed, rather than holding coherent but naïve theories that are consistently applied (Elby, 2000; Smith, diSessa, & Roschelle, 1993). Learning, in this "fine-grained constructivist" view (Elby, 2000), consists of knowing when to apply what resource, and of connecting multiple resources into more complete, coherent, and scientifically normative theories. Student ideas are thus seen as potentially productive steps on the way to mature understanding, rather than as "misconceptions" to be replaced. In this view, resources are not intrinsically correct or incorrect, but are activated for an appropriate or inappropriate context (Hammer et al., 2005). Some resources for graphing include *what you see is what you get* (WYSIWYG – an overly literal interpretation of a graph), *stillness* (the lack of motion suggested by a horizontal line on a graph), *constancy* (the idea that something does not change, prompted by a straight line on a graph), and *sudden change* (suggested by inclined segments on a graph) (Elby, 2000). Not yet studied is how students approach constructing scales, and how the resources they employ can then point to instructional strategies that support deeper learning. The following research question guides our study: What resources do undergraduates employ in constructing scales?

Methods

Our participants are 64 undergraduates at a major public research university in the Midwestern USA, enrolled in an interdisciplinary Geoscience/History course. The students were a cross sample of the university both by year (10% freshmen, 29% sophomores, 25% juniors, 35% seniors) and by major (19% STEM, 52% social sciences, 13% other, 15% undeclared)(Delgado, 2014). All enrolled students participated in the study. We expected that the diversity by year and major would provide insight into a broad array of students' resources.

Students completed identical paper-and-pencil instruments during the first and next-to-last discussion sections. The tasks analyzed here involved constructing one scale for the age of six events and one for the size of six objects, along a 25-cm horizontal line, so that their positions represented their magnitude. The 20 minutes allotted allowed all students to finish the assessment. The time data points were given as years ago (ya), and included: Big Bang (14 billion ya), emergence of life on Earth (3.8 billion ya), emergence of hominids (7 million ya), emergence of Homo Sapiens (600,000 ya), origins of writing (6000 ya), and first human moon landing (at the time, 39 ya). The size data points were given in nanometers (nm) and meters (m) and included: diameter of an atom (0.1 nm), diameter of an adenovirus (100 nm), width of a human hair (0.0001 m), diameter of a dime (0.018 m), height of an elephant (3 m), and length of a football field (100 m). The large range of values was intended to preclude the routine construction of a linear scale, in order to better reveal student resources. Six students were interviewed after completing the tasks for the second time; they were shown their scales and asked to discuss them. The students interviewed were a convenience sample - those who volunteered.

Coding for identification of resources was done using a constant comparison method (Strauss & Corbin, 1998). The data were analyzed in a series of iterative stages. After student interviews were transcribed, both authors independently analyzed them using open coding to identify emergent themes in each student's stated approach to the scale construction task, while also scrutinizing the interviewed students' scales. We identified many common themes among the transcripts and respective scales and came to a consensus by discussion. The interviews and scales then went through another round of coding to refine the resource categories. Both authors discussed any differences among the categories and reached consensus as to their definition and classification. In the final analytic stage, one author developed narrative descriptions for each participant's use of resources, while the other author searched for representative quotations from the transcripts. We then studied the scales of the students who were not interviewed in order to infer other potential resources, using an iterative coding process similar to the one for the student interviews. The findings from scales alone (without interview data) are more inferential and will require additional research to confirm or refine them.

Results and Discussion

Student #11: Show All Elements

This student's four scales were similar. She said they were linear scales, but she did not follow the convention of marking evenly-spaced tick marks and labeling them with numerical values. On a conventional linear scale, the four smallest data points should be indistinguishably close to the origin. She instead spaced the four data points out (see Fig. 1). She said, "So I realized that I probably might have been spacing it out just for space for writing", because "you can only go so small", even though she knew that the data points "should all be at that point right there". She felt the need to clearly portray all data points, even at the cost of accuracy. We call this the *show all elements* resource. It may stem from drawing, where one portrays every important element and usually avoids drawing one object on top of another. This resource is problematic in this case but useful in others. For instance, selecting a range for a scale that just encompasses the data (e.g., 0-25 for data ranging from 2-24) permits the visualization of finer differences between data points than a larger range (e.g., 0-100 or 0-1000). The appropriateness of the *show all elements* resource thus depends on context.

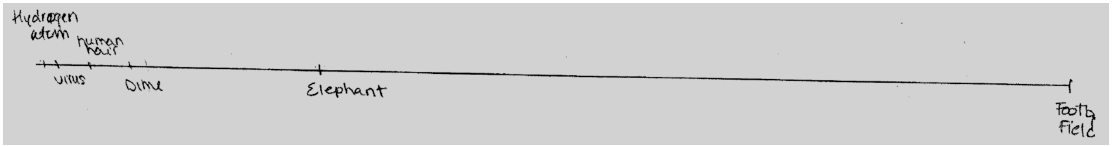


Figure 1. Student #11's scale for size, beginning of course.

Student #46: Linear Scale, Zoom In, Ratio, Conversion

This student activated different resources for time and size. He explicitly stated that his scale for time was linear. His scale conformed to the convention of using evenly-spaced intervals to show equal numbers of units (Leinhardt et al., 1990) - see Figure 2. The conventional linear scale can become a resource in and of itself, composed initially of simpler resources that are coherent and useful for a given context. Over time and through repeated successful use, they become coordinated and constitute a resource as a whole (Hammer et al., 2005). The *linear scale* resource is neither correct nor incorrect *per se* (Hammer et al., 2005). The student realized the limitation of the linear scale for this data set: "I realized I would never be able to cram…". He created a second,

coordinated scale, explaining, "This is me trying to compensate for the fact that I didn't have any space for one billion years". Yet, the extra scale did not solve the problem: "I realized I needed to zoom in much, much further because I still needed to get to 7 million years… If I wanted to do it on a linear scale, yeah, I would have had to keep zooming and zooming and zooming". We term this resource *zoom in*. It may stem from experiences with microscopes, cameras, and adjusting the view in navigation, word processing, and multimedia presentation software. Zooming in is useful for this data set, but would be superfluous with small ranges of data.

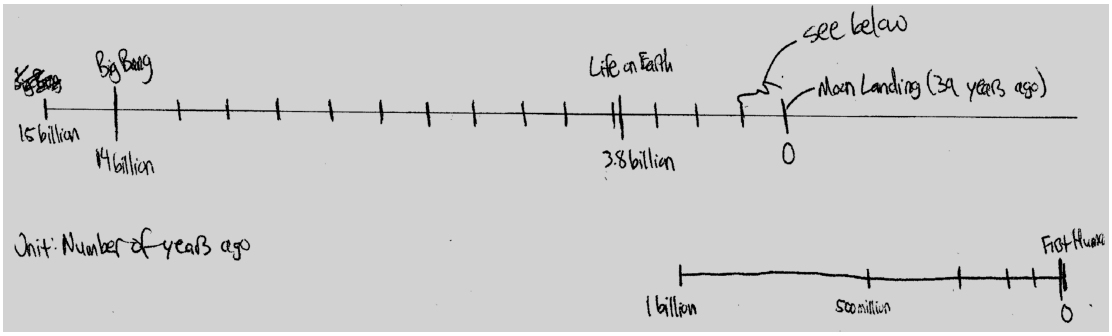


Figure 2. Student #46's scale for time, beginning of course.

The student instead used a *ratio* resource for the size data, focusing on relative size differences between adjacent data points. This approach should in principle result in a scale with spacing identical to a logarithmic scale, although in this student's scale the hair-dime difference is overly large (see Fig. 3). He said:

> I think when I was doing this, I was thinking this [atom-virus] is 1,000 times bigger and this [virus-hair] is another 1,000 times bigger. I left a bigger gap here [hair-dime] trying to compensate the bigger gap. See this jump here [elephant-football field]? This is only 30 times. In reality it should be much much smaller, but I was having trouble illustrating that.

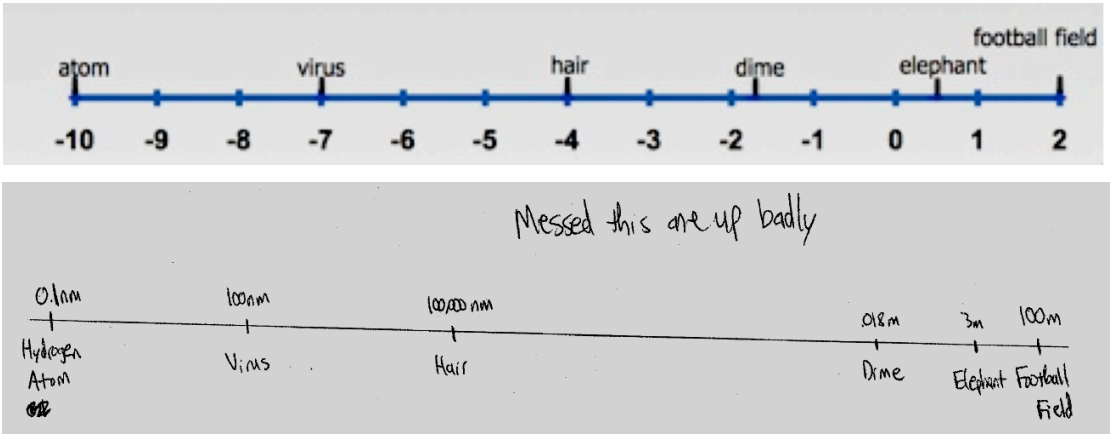


Figure 3. Student #46's scale for size, beginning of course (bottom), and conventional logarithmic scale (top).

This student used nanometers for the smaller objects and meters for the larger ones (see Fig. 3). He noted that "It was probably easier for me to grasp this as 100 nanometers than… 0.0001 [sic] meters… It's all about putting it in terms, for me, that are easy to understand." We identified this as a *conversion* resource. This resource may stem from academic experiences with unit conversions within or across systems, from everyday experiences purchasing food (e.g., ounce to pound conversions) or thinking about time (e.g., changing 120 minutes to 2 hours), or both. This resource is useful for large ranges, but not for small ranges of data.

A striking feature about student #46's scales is his activation of a linear resource for the time data but a ratio resource for size, despite the fact that both data sets had similar ranges and posed similar difficulties: "I wasn't thinking of time as an idea of relative scale. I was thinking of time specifically as a linear, like you would see on a graph… or a time plot…I never would have considered doing time as a relative thing."

Student #17: Powers of Ten, Ratio

This student drew a conventional powers-of-ten scale for size, at the beginning of the course (see Figure 4). Just like with the conventional linear scale, a *powers-of-ten* scale can be a resource in and of itself. However, the student had serious misgivings about his graph: "I remember when I did this it didn't fully represent how… It's just misleading. I mean, the numbers are right, it just makes it that if someone who doesn't understand the

logarithmic of it… Its just going to look like these are, I don't know, the same." He instead drew a proportional scale drawing on the *ratio* resource on the end-of-course scale. This student's unease with a scale that followed conventions for a power of ten scale suggests that some students who learn to draw these may not fully understand them. Supporting the use of a ratio resource – which some students activated spontaneously – may improve instruction of power of ten scales so that deeper understanding is engendered.

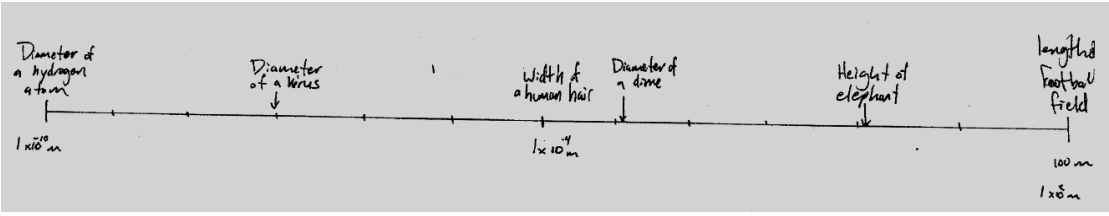


Figure 4. Student #17's scale for size, beginning of course

Student #34: Powers of Ten, Conversion

This student produced conventional powers-of-ten scales explicitly labeled as logarithmic for all four tasks. On the end-of-course scale for size, he used the SI prefixes nano-, micro-, milli-, and kilo-, in combination with the numbers 1, 10, and 100 only (and 0.1 nm, at the smallest end of the scale). Thus, this student used the *conversion* and *powers of ten* resources. Unlike student #17, he had few misgivings about the logarithmic scale. However, like student #17, he also noted that size was easier to think about in terms of proportion than time.

Swarat and colleagues (Swarat, Light, Park, & Drane, 2011) consider that the use of different units for the large and small ends of a scale are indicative of a "fragmented" scale, the least advanced type of scale they characterized. However, thinking of a large range of size data in terms of various units, or even inventing new units – "unitizing" - may be a powerful and appropriate strategy, one used by experts (Tretter, Jones, Andre, Negishi, & Minogue, 2006). The resources of *conversion* and *powers of ten* are useful in this context but would be cumbersome and of little value for small data ranges, which can easily be represented with a linear scale.

Student #48: Halving, Show all Elements

Student #48 produced linear scales for all four tasks. However, he did not use evenly-spaced tick marks to define intervals (see Figure 5). On the first scales, he had no tick marks at all (and yet the spacing of data corresponded closely to a linear scale). On the end of course scales, he used *halving* iteratively to construct his scale: "I tried to put zero at one end and the biggest thing that we had… and just tried to divide it up into similar units. … I'd go down the middle and go like 'that's half, so that's 50, 25, 12.5, 6.'" According to Confrey et al., halving and doubling are "operation primitives", or fundamental understandings (Confrey, Maloney, Nguyen, Mojica, & Myers, 2009). The scale was homogeneous even though the tick marks were unevenly spaced. This student also felt the need to *show all elements*: "I tried to distinguish which ones I thought were bigger and which ones smaller. I don't know if their actual position is represented."

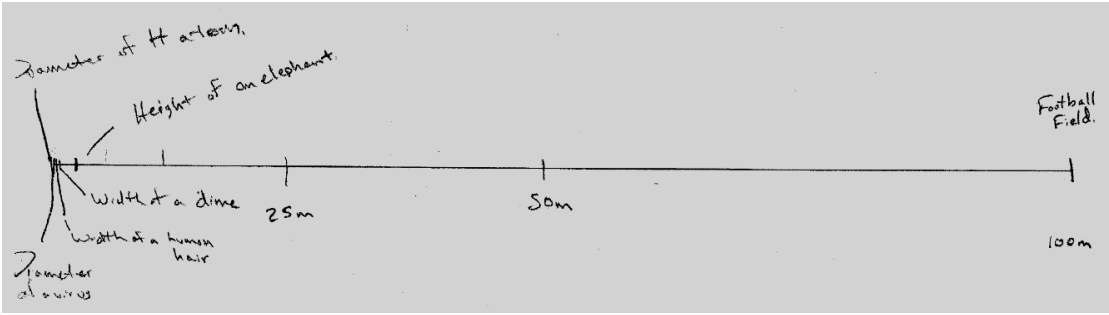


Figure 5. Student #48's scale for size, end of course.

Student #63: Ratio, Logarithmic Scale, WYSIWYG

This student created all four scales employing a logarithmic resource but using "ballpark" estimations with rounded numbers. However, she found the logarithmic scale troubling despite having used it in a physics class:

> Student: Oh for some reason I thought you had to make it what do you call it… Spatially accurate, like to scale. So you'd have this one actually be 100 times longer than this one.
> Interviewer: That would be a linear scale.
> Student: So you're saying a logarithmic scale doesn't even take in to account whether or not it's like, spatially accurate? ... Yeah that's how the physics graph worked 'cause we got rid of the ten and just used the exponents. But then you get a really distorted graph.

She expected that length along the scale would be proportional to the size represented, which is a characteristic of linear scales and also of drawings. In interpreting her own scale, she activated the WYSIWYG resource and was dissatisfied when her scale did not display the transparency she desired. The tendency to represent space as space is a "boundary" that students need to cross in order to produce novel representations and to understand conventional ones (Sherin, 2000). The logarithmic scale does not represent space in a straightforward manner. The teaching of logarithmic graphs in school will need to acknowledge students' preference for and experience with more naturalistic linear representations, and help students contextualize when each is more useful.

Resources Inferred From Scales of Other Students

In this section, we present other possible resources that we infer from scales alone. These are more inferential than the ones identified from interviews and will require additional research to confirm and better understand.

Qualitative Scales or Ordering

Prior research has reported students placing data points at successive evenly-spaced tick marks without regards to the values (e.g., Brasell, 1990), and we encountered some – see Figure 6. This may stem from a *qualitative scale* resource (as learned from bar graphs) for quantitative data. The qualitative scale resource is appropriate for ordinal- or nominal-level data in a bar graph, but sacrifices the relative size information of the quantitative data. It may be that *ordering* is the resource activated here, rather than resources related to bar graphs.

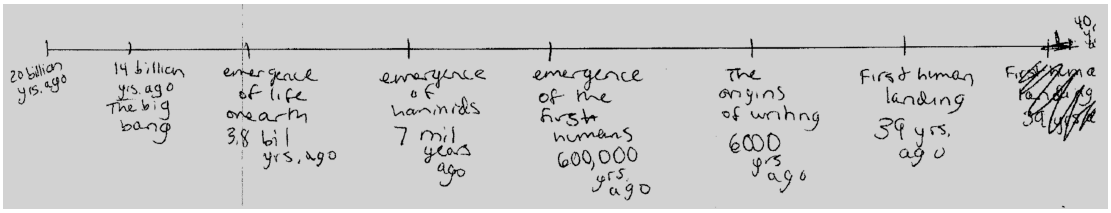


Figure 6. Qualitative scale (Student #1, beginning of course)

Bins

One student created a scale with *bins* or intervals (e.g., 100-1000 years ago). This scale is appropriate for histograms; however, the actual value of quantitative data points is lost. See Figure 7.

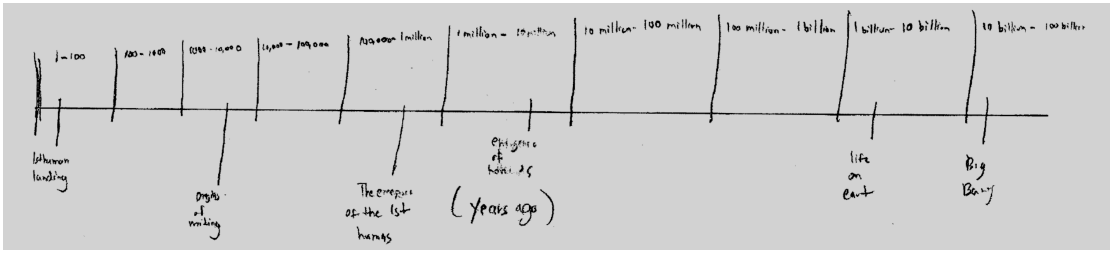


Figure 7. A scale using bins (intervals) (Student #18, end of course)

Scale Break

Some students used *scale break* symbols to indicate that their scales did not include the full range of values. See Figure 8. While the tick marks were evenly spaced in this scale, it is not a qualitative scale because the scale breaks imply a break in the values, which are quantitative. While this resource does not actually solve the problem of representing a wide range of values, it does signal that problem explicitly.

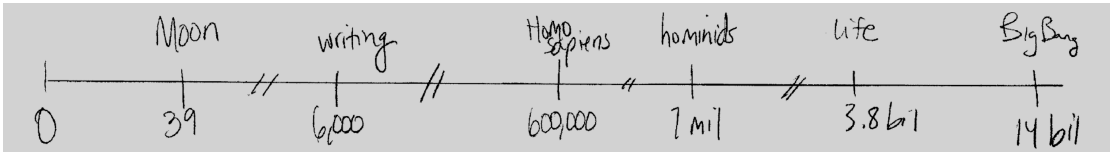


Figure 8. A scale with scale breaks (Student #10, beginning of course).

Writing

One student created a scale with no tick marks. The scale included the name of each event followed by its age, in a sequential manner. See Figure 9. We speculate that the resource he activated was *writing*. This resource would be appropriate (with added punctuation) in a prose account of the data, but not to construct a scale.

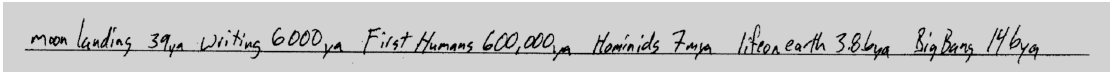


Figure 9. A scale that may reflect the use of a writing resource (Student #29, beginning of course).

Grouping

Some students had an idiosyncratic placement of the data points, with the elephant alone in the middle of the graph. See Figure 10. This "fragmented" scale (Swarat et al., 2011) may stem from a *grouping* resource. The groups might correspond to objects many times smaller than a human, many times larger, and roughly human size, as the human body is a fundamental landmark for size (Tretter et al., 2006). Grouping or classifying is an important conceptual resource, but results in the loss of much information when building a quantitative scale.

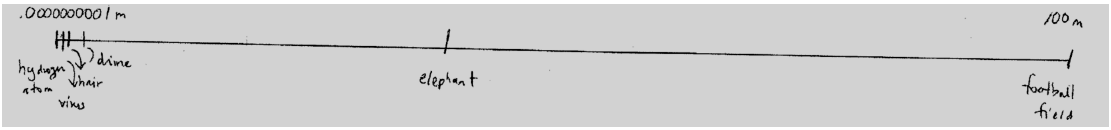


Figure 10. A scale that may reflect a grouping resource (Student #7, beginning of course).

General Discussion

We can conceptually divide the resources we identified into several classes. Drawing has already been shown to be an important source of resources (diSessa, 2004); the *WYSIWYG* (Elby, 2000) resource identified here for scales, and the *show all elements* resources likely derive from drawing. Experiences with graphing also provide resources, some appropriate for this data (*logarithmic, powers of ten*), and some less appropriate or useful (*bins*, *qualitative scales*, *linear scales*). The next class of consists of resources that are appropriate for making sense of the wide range of data, but that would be cumbersome and superfluous with smaller ranges of data: *zoom in, ratio, conversion,* and *scale breaks*. Additionally, some resources employed fundamental knowledge: the operation primitive of *halving* (Confrey et al., 2009), *writing, grouping,* and possibly *ordering*.

The resources students employ, even if they lead to unconventional or suboptimal scales, are ideas on which to build. Additionally, the approximations to conventional representations that students invent, based on their resources, reveal potential pathways of learning. In the next section we discuss specific educational implications of our findings, which extend as far back as elementary school. Further research will be required to study the impact of these suggestions. Limitations of this study include the small number of students interviewed, the reduced racial or ethnic diversity (most students are non-Hispanic Whites), and the fact that the sample was filtered by ability (as these were students at a highly selective university).

Implications

Drawing, Measuring, and Homogeneous Spaces

Even among undergraduates, creating a homogeneous scale was far from universal. Instruction thus needs to create a greater awareness of homogeneous spaces. Drawing might be a useful resource because drawings of a plane orthogonal to the viewer (e.g., a bird's eye view) are homogeneous (if one disregards perspective). The idea that equal lengths represent an equal number of units is used in an implicit manner in drawing; instructional activities could make it explicit through the use of scale factors, beginning with simple ones such as a 12:1 scale where a distance of one foot in the real world is portrayed as one inch on the drawing. Such activities might also build understanding of conversion and ratios. Another early experience with homogeneous scales is the use of rulers. Activities have been developed in which students use collaborative exploration to reinvent measurement, in the elementary grades; these involve reflecting about the constancy of unit size, the need to place units end-to-end without gaps or overlaps, and the idea that measurements do not always consist of a whole number of units (Lehrer, 2003). Such activities may provide resources for graphing quantitative data.

Different Types of Scales and Meta-Level Knowledge of Graphing

Students need to learn to construct and interpret linear, logarithmic, powers of ten, histogram, and qualitative scales, but they also need to acquire the meta-level knowledge surrounding each scale. Instruction may not typically be building this type of knowledge. Students may learn how to graph without linking to their "metarepresentational competence" - their existing understandings of what representations are for, how to create them, and how to evaluate them critically (diSessa & Sherin, 2000). They may be able to produce a fairly conventional scale, yet not truly understand what they have produced. One meta-level understanding is that the choice of representation is impacted by the use it will be given (diSessa et al., 1991). For our data sets, a linear representation might be effective in showing that the Big Bang occurred an extremely long time ago relative to human experience, but is less effective in showing the large relative differences among the more recent events. Logarithmic scales have pros and cons as well: they allow all six data points to be effectively represented, but placing data points that are not exact powers of ten is difficult, there is no way to place zero on the scale, etc. Scales produced using bins, grouping, ordering, or a qualitative approach likewise have advantages and disadvantages. When each type of graph is introduced, meta-level discussion about its idiosyncrasies, conventions, affordances, and constraints is essential so that students can consider when each is appropriate.

Resources for Large Ranges of Data

Students tapped many resources to work with the vast ranges of time and space: through analogy with instruments and software (zoom in), by using proportional reasoning (ratio), by using more familiar numerals (conversion), or by sidestepping the complications of constructing a homogenous scale (scale breaks). The use of similar open-ended tasks with younger students may reveal additional resources that help them make sense of conventional scales and thus graphing. The use of more than one unit on a single scale is unconventional but might scaffold understanding when using large ranges of values. Units can serve as a powerful cognitive tool (Delgado, 2010). However, it would be important to then advance to normative, single-unit scales.

Show all Elements, Ratio, and Halving to Scaffold Learning of Logarithmic Scales

For students to deeply engage in thinking about complex and counterintuitive logarithmic scales, a need to know must be established. The show all elements resource, along with a data set that spans many orders of magnitude, can establish a need to know, since a linear scale fails to distinguish among smaller data points. Instruction of logarithmic scales that builds on students' resources may lead to better understanding than students #17 and #63 displayed. Several students spontaneously used iterative halving to generate unconventional but homogeneous linear scales. This suggests the possibility of systematically using this resource for both linear *and* logarithmic scales. Linear halving (see figure 5) could be paired with multiplicative halving in which *evenly-spaced* tick marks are labeled with iteratively halved data. This would invite discussion of the nature of additive and multiplicative reasoning, scaffolding an understanding of linear vs. logarithmic scales. Since halving is thought to be an operation primitive (Confrey et al., 2009), base-2 logarithmic scales might provide an intuitive way to build conceptual understanding of base-ten logarithms. Other students used a ratio resource. Activities where students develop their own scales based on ratios between adjacent data points (as students #17, 46, and 63 did) could build conceptual understanding of the logarithmic scale. Spatial scale is better suited to ratio reasoning than temporal scale, according to our participants, so size data could be used before time data. A data set with data points that are exact powers of ten and differ by either one or two powers of ten might be useful. Once students realize that 100-fold steps should all be represented by the same length, as should 10-fold steps by a single (different) length, they can explore the size of the two steps. The crucial insight that a 100-fold difference should occupy twice the length as a 10-fold difference could be catalyzed by adding a data point at the geometric midpoint of points with a 100-fold difference. Another potentially powerful insight about the logarithmic scale is that every successive tick mark covers 90% of the remaining range. This idea could be scaffolded by iterative halving, where each tick mark covers half of the remaining range.

Zooming and Logarithmic Scales

Student #46 zoomed into the last billion years of his time scale, but realized he needed to zoom in again and again. Repeatedly zooming in to the smallest *10%* of each scale, and making each scale the same length, could lay the foundations for understanding the logarithmic scale, where each smaller interval covers 10% of the previous. The course textbook (Christian, 2005) featured a series of scales, but they appeared at the beginning of each section rather than together. Juxtaposing successive magnifications may be more effective.

Powers of Ten, Logarithmic Scales, and Homogeneous Spaces

Powers-of-ten scales are not homogeneous. For instance, between 1 and 10 there is a difference of 9 and between 10 and 100 a difference of 90, but each difference is represented by the same distance along the scale. A logarithmic scale transforms those values to 0, 1, and 2, constituting a homogenous space. However, the placement of data points is identical in both. Comparing and contrasting logarithmic and powers of ten scales in high school science or mathematics classes would allow an opportunity to reflect on homogeneous spaces.

Conclusion

This study identifies some resources that are activated when students are tasked with constructing a scale. These resources are cued according to context – for instance, the time data was less likely to activate the ratio resource than size data. The power of the fine-grained constructivist perspective is that unconventional student ideas are not seen as "misconceptions" to be replaced, but as the application of a resource that is useful to students in a different setting. How teachers respond to student ideas is essential. Rather than dismissing them as wrong or seeing them as manifestations of a naïve theory, teachers can use students' ideas to help them build better-organized, more broadly-applicable ideas. As teachers become more aware of these resources, they can better plan instructional activities that allow students to construct more sophisticated understandings and explanations. If we first acknowledge students' resources for representation and build upon them in the classroom, they should develop a deeper understanding of the powerful and diverse standardized representations scientists and mathematicians have developed to represent magnitude and relationships between variables.

References

Brasell, H. M. (1990). Graphs, graphing, and graphers. In M. B. Rowe (Ed.), *The process of knowing*. Washington, DC: National Science Teachers Association.

Christian, D. (2005). *Maps of time: An introduction to big history*. Berkeley: University of California Press.

Confrey, J., Maloney, A., Nguyen, K., Mojica, G., & Myers, M. (2009). Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories. In *PME Conference Proceedings* (Vol. 2, pp. 345–352). Thessaloniki, Greece.

Delgado, C. (2014). Collective landmarks for deep time: A new tool for evolution. *Journal of Biological Education*. Advance online publication. doi: 10.1080/00219266.2013.849280.

Delgado, C. (2010). Units of length: a notational system for conceptual understanding of size and scale. In *Proceedings of the 9th International Conference of the Learning Sciences - Volume 2* (pp. 362–363). Chicago, Illinois: International Society of the Learning Sciences.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition & Instruction*, *22*(3), 293–331.

diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10*(2), 117–160.

diSessa, A. A., & Sherin, B. (2000). Meta-representation: An introduction. *Journal of Mathematical Behavior*, *19*(4), 385–398.

Elby, A. (2000). What students' learning of representations tells us about constructivism. *Journal of Mathematical Behavior*, *19*(4), 481–502.

Hammer, D., Elby, A., Scherr, R. A., & Redish, E. F. (2005). Resources, framing , and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–119). Greenwich, CT: Information Age Publishing.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society.* Cambridge, MA: Harvard University Press.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, & D. Schrifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 179–192). Reston, VA: National Council of Teachers of Mathematics.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1–64.

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). London: Routledge.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and Council of Chief State School Officers.

Nemirovsky, R., & Tierney, C. (2001). Children creating ways to represent changing situations: On the development of homogeneous spaces. *Education Studies in Mathematics*, *45*(1), 67–102.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

Sherin, B. (2000). How students invent representations of motion: A genetic account. *Journal of Mathematical Behavior*, *19*(4), 399–441.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*(2), 115–163.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Procedures and techniques for developing grounded theory.* (2nd ed.). Thousand Oaks, CA: Sage.

Swarat, S., Light, G., Park, E. J., & Drane, D. (2011). A typology of undergraduate students' conceptions of size and scale: Identifying and characterizing conceptual variation. *Journal of Research in Science Teaching*, *48*(5), 512–533. doi:10.1002/tea.20403

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching*, *43*(3), 282–319. doi:10.1002/tea.20123

Acknowledgments

This study was supported by University of Michigan Rackham Merit and School of Education Scholar's Award fellowships to the first author. We wish to express our thanks to the instructor and students of the course. Our appreciation to Leema Berland, Jill Marshall, Brian Fortney, and anonymous reviewers for very helpful comments on this and earlier drafts of this paper.

Scientific Practices Through Students' Eyes: How Sixth Grade Students Enact and Describe Purposes for Scientific Modeling Activities Over Time

Christina Krist and Brian J. Reiser, Northwestern University, Evanston, IL
Email: ckrist@u.northwestern.edu, reiser@northwestern.edu

Abstract: Recent reforms in science education emphasize engaging students in *scientific practices* (NRC, 2011). These reforms aim to not only have students *doing* things scientists do, but to have them doing them *with the similar goal* of constructing explanatory accounts of the natural world in principled, consistent ways. In this case study, we used a communities-of-practice framework to analyze how students' perceptions of the epistemological purposes of several classroom activities change over time. We found that students used their everyday experiences in ways that allowed them to engage in and describe modeling practices that contributed to their classroom's knowledge building goal. In addition, we found that students' articulations of that goal became more epistemologically sophisticated over time. Our analysis provides insights on how students productively use everyday experiences in scientific practices and offers suggestions for how to rethink learning progressions to account for students' perceptions of their modeling practices.

Recent reforms in science education emphasize scientific practices as the means by which students develop scientific ideas (NRC, 2011; Achieve, Inc., 2013). These practices, such as constructing scientific explanations, arguing from evidence, and developing models, are the ways in which scientists build knowledge about the natural world. Thus, the goal in engaging students in scientific practices is to have students doing things that scientists do *driven by a similar epistemological purpose:* to construct explanatory accounts of the natural world in principled ways. As such, classroom scientific practices must connect classroom activities to larger science ideas and principles in ways that help students to make progress in constructing larger scientific ideas themselves *and* in understanding the principled ways in which those ideas were constructed (Duschl, Schweingruber, & Shouse, 2007; Sandoval & Reiser, 2004).

Despite the best efforts of curriculum designers and teachers, students will not be engaged in scientific practices unless they see their activity as *meaningful for their knowledge building*—that is, as connected to their classroom's epistemological goal (Barron, et al., 1998; Duschl, et al., 2007; Sandoval, 2005). However, engaging students in building scientific knowledge requires that most classrooms make significant shifts in how both teachers and students think about the work they do. These shifts, transforming classrooms from places where teachers communicate the ideas of science to students to places where students and teacher work together to build those ideas through scientific practices, take time. As such, classrooms are designed communities that are in the process of *developing* shared epistemological goals and related practices.

Because we want to know how students come to see their classroom activities as meaningful practices rather than routines, we investigated *how students' enactments and perceptions of the epistemological purposes of classroom activities developed over the course of a unit.* In particular, we focused on activities designed to engage students in the practice of developing and using scientific models. In this case study of a classroom working to establish scientific knowledge building practices, we found that students' engagement in and descriptions of modeling activities shifted from using and describing diagrammatic models as *displays of ideas* to using and describing diagrammatic models as *tools for working out ideas*. In addition, students began to recognize peers' roles in working towards their classroom's epistemological goal.

Participation in Scientific Practices in Classroom Communities

The call to engage students in the practices of scientists is not new. However, engaging students in activities that meaningfully contribute to scientific knowledge building is difficult. Hands-on investigations and labs, if not connected to a larger knowledge building goal, do little more than teach students the immediate practical skills necessary for the routine (Barron, et al., 1998). In other words, they gain neither deep content understanding nor a justification for doing the "steps" in the first place. So how do teachers and students work to develop knowledge-building goals and engage in practices that meaningfully contribute to those goals?

Studies of classrooms in which researchers and teachers carefully designed the substance and structure of the context to engage students in scientific practices have found that students successfully engaged in knowledge-building practices when the discourse framing and supporting inquiry emphasized the goal of developing shared knowledge (Herrenkohl, 2006; Rosenberg, Hammer, & Phelan, 2006; Schwarz, et al., 2009). In addition, students developed rich understandings of the incremental building of evidence-based scientific explanations during *sustained engagement* (6 years) with teaching designed to support students' epistemological

thinking (Smith, et al., 2000), suggesting that sophisticated epistemologies take time to develop. Meaningful engagement in scientific practices, then, requires the careful design of sustained sequences of inquiry activities around rich, appropriate content learning goals that are enacted in ways that offer appropriate social supports. Guided by these social supports, students can become active participants in a community that builds knowledge using the tools, social interactions, constructs, epistemological criteria, and discourses of disciplinary science (Ford, 2008). Therefore, viewing science learning as *participation in a classroom community of practice* offers a useful analytical framework for understanding how teachers and students develop knowledge-building goals and learn to engage in meaningful scientific practices.

To examine how classroom communities engage in scientific knowledge building, we interpreted principles about how learning occurs in communities of practice through a lens of epistemological development. In particular, we focused on the *practical epistemologies* that can actively guide the classroom work (Sandoval, 2005), and how those epistemologies develop, or how changes in activation of coherent sets of epistemological resources occur over time (Hammer & Elby, 2002). In this paper, we focused on the characterizing how students' perceptions of *models and modeling activities* changed over time. We anticipated that students could be using models in a variety of ways: to provide a right answer to the teacher; to record his or her own thinking; to explain how and why a phenomenon occurred; or to argue against a competing model (Schwarz, et al., 2009).

However, epistemological development does not happen in isolation. We also wanted to understand how teachers and students used social interactions and shared resources to construct meaning around the modeling work they were doing. We utilized Wenger's (1998) three elements that characterize communities of practice to capture classroom epistemological development holistically. First, a community must establish a *joint enterprise*, or a collectively developed understanding about what their community is about or what they aim to do (Wenger, 1998). Through our epistemological lens, a student will ideally come to understand the epistemological joint enterprise of classroom science—constructing explanatory accounts of the natural world. In other words, students will shift from perceiving inquiry activities as standalone (Barron, et al., 1998), or as serving the purpose of simply propagating knowledge (Hammer & Elby, 2002), to seeing the activity as a means for them to build a larger causal or explanatory story over time (Rosenberg, et al., 2006).

The other two critical elements support the community's joint enterprise. Communities must establish *mutuality,* or norms of interaction and relationship for achieving the joint enterprise (Wenger, 1998). Epistemologically speaking, students will ideally learn to take on important knowledge-building roles that differ from roles afforded by traditional teacher-centered classrooms (Berland & Reiser, 2009). These new roles require that students develop enough trust to be able to both construct and critique each other's knowledge claims (Ford & Ferman, 2006). In addition, communities must support members in using *shared repertoires of resources* in order to accomplish the joint enterprise (Wenger, 1998). In a practices-centered classroom, some of these resources include the discourses, tools, and shared collections of knowledge of scientists. For the purposes of this paper, we focused on a particular set of tangible external resources and their related epistemological resources: those related to the construction and use of diagrammatic models. Students will ideally learn to view diagrammatic models as tools for forming explanatory stories or explanations of phenomenon (Hammer & Elby, 2002; Passmore & Stewart, 2002; Rosenberg, et al., 2006; Schwarz, et al., 2009) rather than as check-your-answer activities, or accumulations and repositories of propagated ideas (Hammer & Elby, 2002).

Based on our framework, as we investigated how students' enactments and perceptions of the epistemological purposes of modeling activities developed over the course of a unit, we expected to see students increasingly: 1. Describing a knowledge-building joint enterprise for their classroom; 2. Describing how classmates' ideas are used in constructing knowledge; and 3. Using and describing diagrammatic models as tools for knowledge building. We found that while students did use models for some of the purposes we expected, they also developed their own epistemological purposes for models that drew on their everyday experiences in ways that contributed to their classroom's knowledge-building enterprise. In addition, we found that they highlighted the importance of peer accountability in knowledge building.

Methods

Research Context

In order to study how students' enactments and descriptions of the epistemological purposes of classroom modeling activities developed, we focused on one classroom using a curriculum designed to engage students in scientific practices (Krajcik, McNeill, & Reiser, 2008). We selected a unit that emphasized the practice of developing and using models to build scientific knowledge because it is a challenging practice for teachers to implement (Schwarz, et al., 2009). This 6th grade classroom is located in a high-achieving middle school in a middle-to-upper-middle-class suburb of a large Midwestern city. The teacher, Mr. H, is an experienced classroom teacher who had taught for 14 years at the time of data collection. However, it was only his second year using this particular curriculum. Although his understanding of scientific practices and pedagogical strategies for engaging students in practices were still developing, he felt that his own orientation to and beliefs

about science teaching aligned well with the knowledge-building goals of the curriculum, especially the focus on having students address and challenge their own ideas and questions about the world with evidence from classroom activities (Interview, 3-18-13). Therefore, although he was still learning how to support students in modeling practices, his commitment to engaging students in knowledge building made his classroom a rich context in which to study the development of classroom scientific practices.

Data Collection

The first author observed and video recorded selected lessons in Mr. H's classroom throughout the 2012-2013 school year. The first curricular unit was these students' first introduction to both scientific practices in general and the specific practice of constructing and using scientific models. The data for this paper comes from the second curricular unit, enacted from January-early April 2013. The lessons selected for observation and analysis (Lessons 1, 4, and 6) were those in which students were constructing, presenting, or revising diagrammatic models and therefore had the potential to be activities in which students were engaged in the scientific practice of developing and using scientific models.

In addition to observing and video recording selected lessons, the first author conducted nine semi-structured interviews over the course of the unit with three focus students from this classroom. Focus students were purposefully selected to represent a range of "getting it," based on the classroom teacher's perception, which likely represents some combination of ability, effort, and interest in science. The interviews were designed to elicit students' perceptions on the purposes or goals of specific classroom activities from Lessons 1, 4, 6. These interviews included questions such as, "Why did you draw a model right away at the beginning of the unit?", "Why do you think you presented your models and ask questions about them?", and "What kinds of things were you thinking about when you were drawing [a specific] model?" As a follow-up to each of these questions, we asked if or how the reasons they gave contributed to their learning in order to elicit their perceptions of how each activity was connected to knowledge building. The first set of interviews occurred between Lessons 1 and 4; the second set occurred shortly after Lesson 6; and the third set of interviews occurred after the end of the unit.

Data Analysis

In order to characterize the shifts occurring over the course of the enactment of the unit, we coded transcripts of both the interviews (163 min. in total) and the classroom observation data (370 min. in total) for *epistemological purposes:* rationales for if, how, and why a particular classroom activity contributed to knowledge building. These rationales were relatively straightforward in students' responses to interview questions. In the classroom video, we coded both explicit statements that described a particular epistemological purpose (e.g. "Felix is doing something good here, he's making connections to things we did in the light unit," coded as *Compare to things we've done*) as well as statements that implied or operationalized a particular epistemological purpose (e.g. A student saying, "But there can't be empty space because air has to expand" in response to another student's model, coded as *Compare to things we know*). From the 9 semi-structured interviews, we first generated "in vivo" codes (Miles, Huberman, & Saldana, 2014, p. 74) capturing the epistemological purposes that students described. Drawing on Hammer & Elby's (2002) categories of epistemological resources, we then collapsed our initial list into 19 "epistemological purpose" codes, each representing a particular combination of *epistemological activity + nature/source of knowledge*. These codes included purposes such as, "Make thinking visible," "Public evaluation," and "Shift individual understanding." We then applied these codes both the interview data and the classroom discourse. Applying these codes to the classroom video generated one additional code: *Compare to real life experiences* (see Table 1).

After coding all the data, we eliminated epistemological purpose codes that were not a) included at least twice in the three focus student interviews from the class period of interest, and b) included at least once during the classroom discourse. Although we acknowledge that any classroom activity has multiple overlapping goals and purposes, this reduction was an attempt to focus on the *most salient purposes* at each point of time during the unit as enacted by the teacher and students AND as interpreted by students in interviews.

Characterizing Classroom Enactment and Students' Perceptions of Scientific Practices

We argue that over the course of the unit, students' enactment and descriptions of the purposes of activities related to constructing and using scientific (diagrammatic) models shifted from enacting and describing diagrammatic models as *displays of ideas* to describing and enacting models as *tools for working out ideas*. In the classroom enactment, the class shifted from using their models for displaying or collecting knowledge to using their models as revisable representations of the ideas they were comparing and actively (re)building. This shift in enactment preceded a parallel shift in how students explicitly described the purposes for modeling activities. However, rather than describing models as tools for working out ideas, students instead described the social mechanisms through which they used their models for knowledge-building. Taken together, these shifts

suggest that students were developing meaningful ways of participating in scientific modeling practices. In this paper, we characterize these shifts first by looking at classroom activity, supported with data from student interviews. We then analyze student interviews for additional ways in which students developed meaningful modeling practices. Finally, we discuss the implications of these shifts for epistemological development.

Table 1: Epistemological purpose codes.
(These codes were found at least twice in the three focus student interviews and found at least once during classroom discourse.)

"Epistemological Purpose"	***Epistemological Activity + Nature/Source of Knowledge (from Hammer & Elby, 2002)***	**Paraphrased descriptions of the point of the activity**
Make thinking visible	*Displaying + Inherent*	The point is seeing each other's ideas to get a sense of what students think
Record thinking for future use	*Accumulating + Propagated*	The point is to keep track of what we know for use on a test in the future, to prevent myself from getting confused, or to compare later and see how much we've learned
Compare to things we know	*Comparing + Propagated*	The point is to think about the facts we've learned to see if a model or idea matches with them
Compare to real life experiences	*Comparing + Direct Perception*	The point is to think about things we've seen, experienced, or know from real life and see if a model or idea matches with them
Compare to things we've done	*Comparing + Fabricated and/or Direct Perception*	The point is to think about the experiments and discussions we've had in class to see if a model or idea matches with them
Revise representation through peer feedback	*Formation + Fabricated (specific ideas represented)*	The point is to change my model based on the questions and critiques my peers gave me
Learn through collaboration	*Formation + Inherent (in other) and/or Fabricated*	The point is to use peers or teacher as a resources as they ask questions, give critiques, or offer their ideas and understandings
Solve or piece together puzzle	*Formation + Fabricated (big ideas)*	The point was to make another connection (snap another piece in) that moved us closer to solving the overall puzzle/question

Models as Idea Displays: Seeing Ideas and Making Comparisons

During the first lesson of the unit (Jan 10th-11th, 2013), Mr. H had students smell an odor in a film canister and discuss how they thought the odor moved through the air so they could smell it. Students then drew a model illustrating their ideas about what makes up the odor if they were to magnify it. In framing and enacting this lesson, both Mr. H and the focus students explicitly described the purpose of modeling activities as *Making Thinking Visible* (see Table 1). When students began drawing their models during Lesson 1, the teacher emphasized that students were to be drawing "YOUR ideas, not you and your neighbor." He then said that once enough people had something on their paper, they would "take a peek at" each other's models and "share out" their ideas. After giving students a few minutes to draw, he asked if anyone was ready to share their models, stating, "Don't worry, we're not here to judge, especially since none of us has the answers. We're all just getting started with this, so we're curious. We're curious what you think." In framing this activity, Mr. H explicitly denied any evaluative purpose.

In retrospective interviews following Lesson 1, two focus students described the purpose of drawing and presenting these initial models in ways that mirrored Mr. H's framing of the activity. Carly said the purpose was "just to see like what our preference is on it, like what we think it is." Similarly, Ruthie said the purpose of sharing their models was so that they could see what everyone else was thinking. Again, neither of them described an evaluative goal. In and of itself, the epistemological purpose of "making thinking visible" does not suggest any connections or contributions to a joint enterprise; rather, students are each displaying their own inherent ideas. Here, it seems that Mr. H was working to position this activity in contrast to the evaluative purposes that frame most presentation-like school activities.

As students presented their models, however, they did not simply note each other's ideas. Instead, the class began to make comparisons between the previous unit (the "light unit") and the current one (the "smell unit"). Mr. H first connected the two units by guiding students to think more deeply about how they smelled the odor: "Now I know that you guys are pretty careful observers, and you noticed that smelling whatever is in here, that was actually a sort of process. There's a beginning, a middle, and an end. I remember for our light unit we talked about light starts somewhere, goes to an object, bounces off the object, hits our eyes, and we know there's a lot more too with color now. But there's a process, right?" Mr. H made an explicit connection to how they had been working on ideas in the light unit, drawing a parallel to how they would investigate this new

question about how odors travel. This connection was his first attempt at establishing a joint enterprise: to figure out the process by which an odor moves across a room. Notably, this enterprise is a knowledge-building one.

During their model presentations, students also made connections to the light unit in ways that contributed to the newly established joint enterprise, suggesting they were at least somewhat bought in to Mr. H's initial framing. Felix, who presented first, used ideas from the light unit to highlight how his ideas about how odor moved contrasted with light: "And then you can see like where the odor moves around in all these different directions. So it doesn't have to, unlike with the light it doesn't really have to travel in straight lines either. So it goes anywhere it wants, really." Lola, the second student to present, also highlighted a difference between the how light and odor travel. She represented odor as a line with a curve in it to indicate "the odor is not like straight lines, but like where the wind takes it and curves it around corners."

In response to Lola's presentation, the class began to bring in ideas from their everyday experiences to justify their comparisons. Mr. H asked the class if they agreed that odor could move around corners. One student said, "You can smell something that's on the other side of the house." The class briefly discussed smelling cookies baking, and then Lola turned to Mr. H and said something inaudible to the camera. He turned to the class and shared her question, saying, "All right, now I don't have any answer for this, because her question was […] asking if temperature has anything to do with this, with odors." Several students responded, all seeming to be in agreement that temperature did have to do with smelling odors. Mr. H said he thought they would need to collect some evidence to decide if temperature was a factor. The class agreed, and Lola wrote her question on a Post-It note.

In these two presentations (and the four that followed), students used their models to display ideas that highlighted contrasts between their new ideas about how odor traveled and their knowledge from the previous unit that light travels in straight lines. So they were both "displaying" ideas, but they were also "comparing" to what they all knew. In addition, to support this epistemological activity of comparison, they began bringing in real life experiences—the experience of smelling food around corners—to help them generate potential factors involved in odor travel and provide initial justifications for whether or not those factors mattered. In other words, students were using their models as a visual platform for making comparisons and generating questions based in their everyday experiences. Importantly, the teacher affirmed these moves and valued questions as important products of the discussion: each time a student made a case for a new potential factor based on a personal experience, he gave the student a Post-It note to record the question and made comments like, "I'm curious too!" and "Good, we're getting somewhere." Through his affirmation, Mr. H established the joint enterprise for the unit and acknowledged that the ways students were engaging in the activity—recording and displaying their ideas and making comparisons to generate questions—contributed to that enterprise.

Models as Thinking Tools: Making Sure Our Ideas Make Sense

As the unit progressed, students continued to draw on their prior knowledge and experiences in modeling activities. In Lesson 4 (Jan 30th, 2013), making comparisons to prior experiences allowed students to initiate the one instance of sustained argumentation between competing ideas that occurred during the smell unit.

By this point in the unit, the class had decided that air was an important factor in how an odor moved and were investigating how air behaved when expanded or compressed. Lesson 4 began with Mr. H adding and removing air from a sealed flask. Students then drew what they thought air looked like in a normal sealed flask, in the flask where air was removed, and in the flask where air was added. After working individually for a few minutes, two students presented their models. The first student, Summer, presented a rather elaborate model explaining that when air was removed from the flask, each individual air particle expanded to fill up the space. Likewise, when more air was pumped into the flask, each individual air particle shrank to allow more particles to fit in (see Figure 1). Students asked her a few questions, including what was in between the particles. She stated that there was more air in between; she just did not have time to draw that many particles.

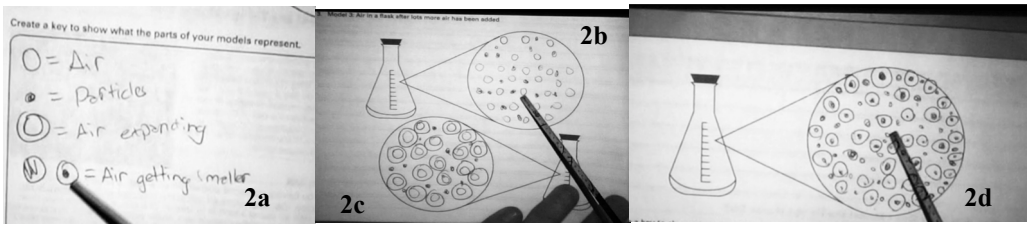


Figure 1. Summer's Models, Lesson 4. 2a: Key. 2b: Normal air. 2c: Air removed. 2d: Air added.

The next student to present, Jared, claimed that there was "nothing" between the particles: when you added air, there was more air and less nothing, and when you removed air, there was less air and more nothing (see Figure 2). Almost immediately, the class erupted into a heated discussion about his idea, arguing that it was not possible to have empty space between particles. Summer stated several times, "there can't be empty space,

because air has to expand" because "otherwise it wouldn't be a gas," highlighting a contrast between Jared's model and a principle they had learned in class. Nate tried to imagine[1] what it would mean for there to be empty space. He asked, "Would that mean if I stepped into it, I would shrink or something? If there was no air?"

Jared then attempted to defend his idea by drawing on a real-life experience: he asked, "What about in space?" Summer, supported by many other students, argued that there is air in space, it is just THIN air because the individual particles have each stretched out so much that they are very thin. Robbie and Dexter both added that air in space is like thin air at high elevations, which is why climbers on Mt. Everest need oxygen. Here, the class co-refined Summer's initial idea to make it fit the knowledge they had about places with "thin air."

Jared then tried another tactic to defend his model: he described what having no air between particles would look like on his model. He said, "I think if [their idea] was true then, I think [the models] would all be like the same, like that [*pointing to his first model*]." Mr. H asked, "They would what?" Jared clarified, "They'd all look like that, they'd all be the same. The same amount of particles there. Cuz they're saying that there's air in here? [*pointing to empty space in his model*]." Unfortunately Mr. H did not understand what Jared was trying to say. He asked if that meant all the lines would get fatter, and Jared gave him a confused look. Then another student asked an off-topic question, and Mr. H adjourned the class for the day.

Despite this unsatisfying end to the argument, we want to highlight that rather than models simply displaying ideas, they now display ideas *for a purpose*. Students considered whether or not they were persuaded by another student's model and expressed a need to resolve the discrepancies they saw. In addition, Jared made a move that was, as of yet, unprecedented in this classroom: he attempted to use his model as a tool to rebut a counterargument. His model made his (and Summer's) ideas visible *in order to* determine which idea better explained this phenomenon, based on what they knew to be true about the world. In this episode, students used models as tools to work through and form ideas together about what air looks like so they could figure out how odors move through it. Importantly, the mechanism by which they worked through and formed these ideas was by drawing on and comparing displayed ideas to their prior knowledge and everyday experiences in the world.

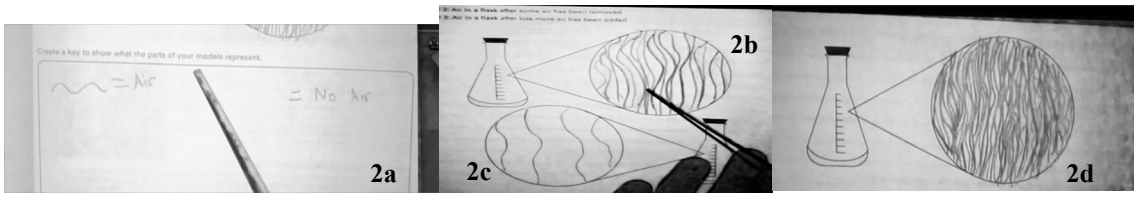


Figure 2. Jared's Models, Lesson 4. 2a: Key. 2b: Normal air. 2c: Air removed. 2d: Air added.

By Lesson 6 (Feb 13th, 2013), the class returned to their original question about how odors move across a room. They saw an animated simulation that showed particles traveling in straight lines until they bounced off of each other. They then worked in small groups to draw a revised model showing what the odor looked like in between the odor source and the nose. In an interview following this activity, Felix explained how he used the activity of drawing his model in Lesson 6 as a tool to help him work through ideas. He described how, as they were drawing, he and his partner wondered about how particles would bounce: "I was also wondering like how it would work, like in what direction would [the particle] go in? Let's say it hits like the exact corner of a wall, directly at the corner, like would it just bounce off, or would it like scatter away, or would it split up? Probably not, but you know I was wondering about that, you know." Here, Felix used the process of drawing the model to help him think through questions he still had and to articulate the specifics of his explanatory account. He later decided that particles would not split apart, and that they needed to hit obstacles placed at different angles in order to spread all around the room. For Felix, the diagrammatic model itself was not the goal; rather, the purpose was to think through his ideas to figure out how the odor was moving

Although Jared and Felix used models as tools for refining ideas in practice, none of the students explicitly stated an idea-refining purpose for models in interviews, such as using models to compare or decide between competing ideas. Instead, of the 15 purposes they did describe, 10 of them were about making comparisons to things they had done or experienced. This is not surprising, given the prevalence of connections they made during Lessons 1 and 4 and the lack of uptake of Jared's use of his model. However, students did describe modeling activities as serving a knowledge-building goal more generally. In doing so, they described how the social aspect of knowledge building helped them to refine their models. Their descriptions suggest that changes in social roles, or establishing *mutuality*, were more salient for students than changes in how they were using models to build knowledge. We turn next to their interviews to illustrate this emphasis.

Describing How Peers Connect Models to the Joint Endeavor

Although students enacted modeling practices in ways that suggested they saw them as meaningfully connected to the joint endeavor, we also wanted to see how students' articulation of the connection developed over time. We found that by midway through the unit (during their second interviews), students explicitly described, in

their own words, the joint endeavor of their classroom in ways that involved both metaphors for building ideas and implied a necessarily social process. When asked about drawing models, Ruthie said, "I like how this class is […] sort of like a mystery that you unravel day by day, so *we're* sort of like invested in trying to figure out like what's the big like secret that *we* always like have (*italics added)*." Carly described model drawing in a similar way, saying that when drawing a model *in a group* you can take in *other people's* ideas and perspectives and "piece together the puzzle." From these statements, we see that these students were beginning to see their modeling activities as tools that allowed them to build knowledge *together*.

Interestingly, students' third interviews at the end of the unit described more articulately how peers influenced their classroom knowledge building during modeling activities. This shift is especially striking, as the interview questions remained the same. Felix described how he considered peers' ideas during model presentations: "I'm paying attention to what I don't have, […] but I would ask like, why would you put this in? And then if it gave me like a good explanation for it, I'd like think about it and try to put it into my [model], if it was like really good." Here, Felix described a three-step accountability process: he notices differences between the presented model and his own, he asks for the presenter's rationale, and he decides if the rationale is good enough or not. Ruthie also articulated a three-step accountability process when she described how discussing someone else's model helped the presenters: "Because if you're like telling them, […] like, you should probably do this, and then they could, and then the whole class could like join in and see if that's a good idea or a bad idea, because sometimes you have bad ideas, and you share them, but then like the class like keeps you in check." According to Ruthie's description, a presenter shares an idea, a student offers a suggestion, and the class decides if that suggestion is worthwhile or not. Although these students are not describing *how* they decide if an idea is good or bad, a process for which science has very explicit criteria, they are recognizing that knowledge building in science class requires accountability for ideas. In other words, not only do they recognize that their models and modeling activities are resources they use to contribute to the joint enterprise, but they are also beginning to recognize that there are particular ways to use those resources, and that using them requires that each member contribute. Their participation in modeling activities is not just another classroom routine. Instead, it is a meaningful practice in which they, as a class, build scientific ideas together.

Implications from Students' Enactments and Perceptions of Practices

We have shown how over the course of the unit, this class gradually developed meaningful ways of engaging with scientific models that went beyond simply adding a new school routine. First, they used models simply to make their thinking visible. However, students quickly took up the new joint enterprise—figuring out the process for how odor moves—and began making comparisons to prior knowledge and experiences in order to generate new ideas and questions. In their next modeling activity, students continued to make comparisons to things they knew or had experienced. They used these comparisons to initiate the one instance of extended argumentation that occurred in this classroom during this unit, using their model displays as tools to work through ideas together. Interviews with students suggested that, both individually and in groups, they continued to use their models to work through, rather than simply display, their ideas, and that peer accountability played a salient role in how their knowledge building worked. These shifts suggest that modeling activities in this classroom were developing as instantiations of a meaningful, purposeful scientific practice.

So how did these meaningful epistemological purposes for modeling activities develop over time? In this case study, two important activities afforded students both epistemic authority and epistemic accountability. First, this class engaged frequently in the epistemological activity of *making comparisons*, especially when the knowledge source was students' *shared or everyday experiences in the world.* These students often engaged in "everyday sensemaking" (Warren, et al., 2001) during scientific modeling activities. In this classroom, everyday sensemaking afforded students the epistemic *authority* to construct knowledge and allowed them to argue against a Jared's claim even without deep content knowledge or scientific expertise. Second, students described how their class engaged in a simple form of classroom *accountability* to keep each other "in check." At least to our focus students, peer accountability was a salient and purposeful part of their knowledge building.

It is important to note that students' forms of everyday sensemaking were valued and even praised by this classroom teacher. The challenge for the teacher here was not learning to make sense of, recognize, and value students' everyday ideas, as was the case in other studies (e.g. Lee, 2001; Warren, et al., 2001). Instead, this teacher struggled to connect students' everyday sensemaking to more disciplinary ways of engaging in those practices—or with balancing students' epistemic authority with *disciplinary* accountability (Ford, 2008), or guidance for students in *how* they decide if an idea is good or bad. Mr. H's difficulty in facilitating argumentation demonstrates the need for explicit guidance or "rules of thumb" for teachers in how to help students engage in more meaningful versions of scientific practices. Learning progressions for practices, then, should help teachers not only in understanding the disciplinary versions of scientific modeling practices but also in how to use moments of sophisticated practice driven by everyday intuitions into more disciplinarily-consistent versions of the practices. In other words, we need tools that help teachers know how to balance epistemic authority and accountability in order to develop students' modeling practices *deeply,* connecting

everyday and disciplinary practices at each stage of a learning progression, in conjunction with tools that help teachers move students from simple to more sophisticated versions of scientific practices.

Endnotes

(1) Although Hammer & Elby (2002) list "imagining" as a distinct epistemological activity, we considered the purpose of Nate's move to be parallel to Summer's: testing to see if Jared's idea was compatible with things they knew or had experienced. Therefore, we coded Nate's utterance as an instance of *Comparing to [imagined] real-life experiences.*

References

Achieve, Inc., on behalf of the twenty-six states and partners that collaborated on the NGSS. (2013). *Next Generation Science Standards*. Achieve, Inc.

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem-and project-based learning. *Journal of the Learning Sciences*, *7*(3-4), 271-311.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404-423.

Ford, M. J., & Forman, E. A. (2006). Redefining disciplinary learning in classroom contexts. *Review of research in education*, *30*, 1-32.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.) *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). Mahway, NJ: Lawrence Earlbaum Associates, Inc.

Herrenkohl, L. R. (2006). Intellectual role taking: Supporting discussion in heterogeneous elementary science classes. *Theory Into Practice*, *45*(1), 47-54.

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning‐goals‐driven design model: Developing curriculum materials that align with national standards and incorporate project‐based pedagogy. *Science Education*, *92*(1), 1-32.

Lee, C. D. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal*, *38*(1), 97-141.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. SAGE Publications, Incorporated.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington DC: National Academies Press.

Passmore, C., & Stewart, J. (2002). A modeling approach to teaching evolutionary biology in high schools. *Journal of Research in Science teaching*, *39*(3), 185-204.

Rosenberg, S., Hammer, D., & Phelan, J. (2006). Multiple epistemological coherences in an eighth-grade discussion of the rock cycle. *The Journal of the Learning Sciences*, *15*(2), 261-292.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634-656.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic supports for science inquiry. *Science Education*, 88, 345– 372.

Smith, C. L., Maclin, D., Houghton, C., & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, *18*(3), 349-422.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632-654.

Warren, B., Ballenger, C., Ogonowski, M., Rosebery, A. S., & Hudicourt‐Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday sense‐making. *Journal of Research in Science Teaching*, *38*(5), 529-552.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

Acknowledgments

This research was funded by the National Science Foundation grant DRL1020316. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors.

The Impact of a Social Robot's Attributions for Success and Failure in a Teachable Agent Framework

Kasia Muldner, Victor Girotto, Cecil Lozano, Winslow Burleson, Erin Walker
Computing, Informatics & Decision Systems Engineering, Arizona State University
{katarzyna.muldner, victor.girotto, cecil.lozano, winslow.burleson, erin.a.walker}@asu.edu

Abstract: Teachable agents foster student learning by employing the *learning by teaching* paradigm. Since social factors influence learning from this paradigm, understanding which social behaviors a teachable agent should embody is an important first step for designing such an agent. Here, we focus on the impact of causal attributions made by a teachable agent. To obtain data on student perceptions of agent attributions, we conducted a study involving students interacting with a social robot that made attributions to *ability* and *effort*, and to the *student*, *itself*, or *both*. We analyzed data from semi-structured interviews to understand how different attributions influence student perceptions, and discuss design opportunities for manipulating these attributions to improve student motivation.

Introduction

Interactive activities can be highly beneficial for learning, because they provide opportunities for knowledge construction through, for instance, contributions to group discussion or guidance from a knowledgeable partner (Chi, 2009). While pedagogical interactions have historically been student-student or student-teacher, as educational technology evolves, an emerging avenue has involved using pedagogical agents to foster learning through agent-student interactions (Woolf et al., 2010). One type of pedagogical agent is a teachable agent, which simulates the collaborative activity of peer tutoring, where it is the student who teaches the agent about the target domain. Prior work has demonstrated that there are cognitive and social benefits to peer tutoring (Roscoe & Chi, 2007), and by extension, to the teachable agent paradigm (Chase et al., 2009). In our work, we are interested in exploring how social interactions between a student and a teachable agent foster student engagement with the agent, and ultimately improve learning and motivational outcomes.

We focus on a specific type of social behavior, namely the *attributions* that a teachable agent makes. In general, attributions correspond to causal explanations one makes for successes or failures. According to attribution theory, the causes that students attribute to outcomes impact their motivation, affect, and reactions, as well as subsequent learning outcomes (Försterling, 1985). Moreover, listening to others' causal attributions influences the overhearing student's affect and social behaviors (Hareli & Weiner, 2002). This prior work, however, has been done in the context of a classroom with students. It is an open question, therefore, as to the impact of a teachable agent that makes causal attributions. We conducted a study gathering data from students' interactions with a social robot that makes different types of attributions to ability and effort. Our general research question was as follows: How do the different types of attributions made by a teachable agent impact students' perceptions of the agent and their desire to teach the agent?

To address this question, we rely on a robotic teachable agent platform that we have developed called TAG (*Tangible Activities for Geometry*) (Muldner et al., 2013). To the best of our knowledge, all related efforts using the teachable agent paradigm have focused on agents in virtual environments. In TAG, students instead walk around a projected space and interact with a physical robot. There may be several advantages to a robotic learning platform. First, a physical presence provided by a robotic agent strengthens users' perceptions of having a social partner more than a virtual agent (Powers et al., 2007). Second, students benefit from learning through embodied, physical interactions, particularly for abstract topics (O'Malley & Fraser, 2004), which robotic platforms naturally support. Thus, in a robotic learning environment, the effects of a robot's social behaviors may be heightened, and so it is a good platform for testing the effects of a teachable robot's attributions. In the remainder of this paper, we present related work on social behaviors in teachable agents, and virtual and robotic agents. Next, we describe the TAG system and its social robot Quinn, and then present a study investigating how students perceive Quinn's attributions. We conclude by identifying implications of the results for designing a teachable robot to engage students and ultimately improve their learning outcomes.

Background

Teachable agent systems have evolved out of personalized learning research and have yielded some early successes. Teaching a computer agent can lead to more learning than being taught by a computer agent (Leelawong & Biswas, 2008) and is nearly as effective as being taught by an expert tutor (Reif & Scott, 1999). Some of the teachable agent effect is due to the deep cognitive processes fostered by teaching: As in peer-to-peer tutoring, peer-to-agent tutors notice their own misconceptions and elaborate on their knowledge as they tutor their teachable agents (Biswas et al., 2005). Another factor responsible for the benefits of learning from

teaching is motivational. For instance, students feel responsible for their agent students, and as a result try harder and attend more to subject material (Chase et al., 2009). To capitalize on these motivational aspects, researchers have begun taking steps to build social and affective behaviors into their agents. For example, Gulz et al. (2011) have incorporated off-task social conversation into their teachable agent, and demonstrated that this led students to learn and have a positive attitude. Others have begun to explore how conversational strategies such as teasing between human peers (Ogan et al., 2012b) and human-agent peers (Ogan et al., 2012a) impacts rapport with the human learner. These efforts are at an early stage, and so more work is needed to understand how to capitalize on social and affective elements within a teachable agent platform.

However, in the broader pedagogical agent literature, we do have some information on the impact of social agents, both in the virtual and physical domains. In general, social behaviors can have a positive impact on student perceptions and in some cases learning. As far as virtual agents are concerned, students preferred agents who display facial expressions over ones that do not (Baylor & Kim, 2008). Others have explored the impact of social behaviors in robots. Kanda et al. (2012) had students interact with either a task-oriented robot or one that also provided social support by praising and encouraging students. Although no learning difference was found between the two versions, students preferred the social robot and reported a stronger relationship with it. Saerbeck et al. (2010) showed that students who interacted with a socially-supportive robot (e.g., one with facial expressions) learned more and were more motivated than students who worked with a neutral robot. Leite et al. (2010) found that users reported higher feelings of companionship with a robot that empathized with them during a game of chess, as compared to a robot that did not.

Another promising way of socially engaging students is through the use of attributions. For instance, virtual agents that express attributions emphasizing the utility of perseverance have been shown to improve students' affect during problem solving (Woolf et al., 2010). The teachable agent paradigm gives us a unique platform for exploring the effects overhearing a teachable agent's attributions. We focus on attributions to effort vs. ability, and attributions to the peer tutor, the robot, or both. Our target set of attributions includes both "desirable" and "undesirable" ones, because we want to explore their impact on student perceptions. Some work outside of computational frameworks has highlighted ways that overhearing a student make attributions impacts the observer (Hareli & Weiner, 2002). For instance, for unsuccessful outcomes, hearing a student attribute the failure to low ability can trigger either pity or contempt in the observer, while attributions to low effort can elicit anger in the observer if they are teaching. For positive outcomes, less is known about how a teaching framing influences observers' attributions, but Hareli and Weiner (2002) speculate, for instance, that attributions to effort are perceived as modest and so can contribute to feelings of admiration in observers. While informative, work is needed to extend and refine these findings to computer environments with teachable agents.

Tangible Activities For Geometry (TAG)

The *Tangible Activities for Geometry* (TAG) system that we use as the test-bed for our work aims to help students learn about geometry by providing a projected space that students move in while solving problems and interacting with a robotic agent called *Quinn*. TAG comprises four main components (see Fig. 1, left): *Quinn*, the *problem space*, the *hanging pointer*, and the *mobile interface*. Quinn consists of a LEGO Mindstorms robot with an iPod mounted on top of it representing its face. The problem space is projected on the floor and includes a Cartesian plane and Quinn, which moves autonomously in this space (for details, see Muldner et al., 2013). The hanging pointer is TAG's version of a mouse. It corresponds to a small cylinder attached to the ceiling by a retractable wire, and is used to control a virtual pointer projected in the problem space (i.e., moving the hanging pointer results in the virtual pointer following it). To click on objects in the projected space, students hover the hanging pointer over the desired target (either a point or Quinn), and pull the pointer down towards the ground and back up. The mobile interface is an iPod Touch that lets the student see the current problem, move between problems, check for correctness of the current solution, and issue commands to Quinn (Quinn responds after each student instruction by executing that instruction).

To illustrate student interaction with TAG, suppose a student opens the problem: "*Plot the point (2, 1)*". When a new problem is opened, Quinn moves to the origin and faces east along the x-axis. To solve the problem, the student must walk over to Quinn and click on it using the hanging pointer. Clicking results in the student's iPod showing the list of available commands to give to Quinn. As the first step, the student could tap *move* and specify the distance '2' in his/her iPod, which results in Quinn moving two units along the x-axis. The student must then walk over to Quinn, click, and choose *turn in a direction*, specifying 'N', which results in Quinn facing North. The last two actions correspond to moving Quinn by 1 unit and then telling it to plot a point. When ready, students can tap a button on their iPod and correctness feedback is shown on that iPod.

We chose the current task domain, i.e., geometry, because of its conceptual properties. In theory, as students move over the projected coordinate system and gesture towards aspects of the projection, they can physically encode concepts such as how positive and negative coordinates relate to graphical quadrants.

Quinn's social behaviors. Quinn's social behaviors are based on attribution theory, and are generated after and in response to TAG's feedback for correctness, ostensibly representing the robot's reaction to whether

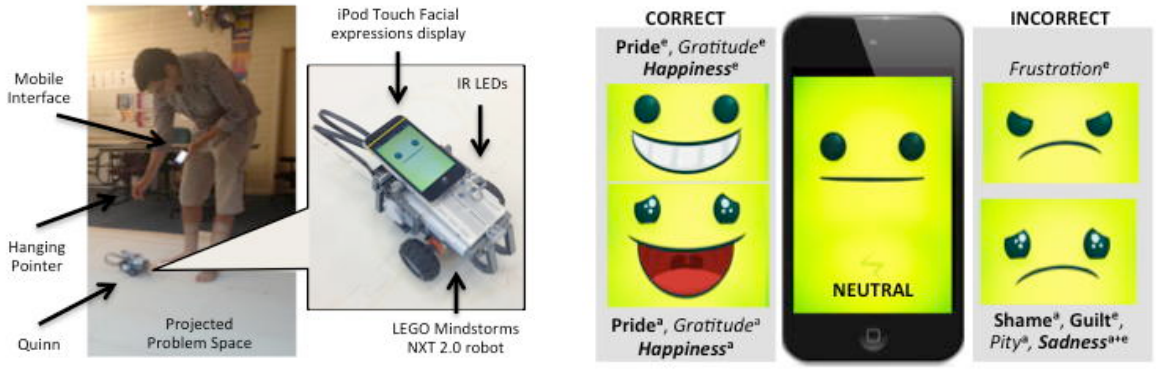


Figure 1: The TAG system (left) and Quinn's facial expressions (right, including text specifying attribution-specific emotions based on the attribution dimension: I (bold), you (italics), we (bold italics); a = ability and e = effort attributions. Students did not see these labels). Note that a single expression is shown at a given time.

it got the correct answer. Specifically, Quinn responds to TAG's feedback by displaying an emotion on its iPod (see Fig. 1, right) and by telling the student how it feels through a message spoken in a gender-neutral voice. In the message, Quinn attributes the outcome to factors along two dimensions: the cause of the outcome (*effort* or *ability*) and the agent responsible for that cause, namely itself (*I)*, the student (*you)*, or both (*we*). Thus, there are six messages for correct outcomes and six for incorrect outcomes (see Table 1). We focus on ability and effort because they are the most common attributions students use to explain academic outcomes (Hareli & Weiner, 2002). Some of Quinn's attributions are undesirable (e.g., attributing failure to lack of ability). We included the full spectrum of messages because we wanted to comprehensively explore the impact of various attributions on student perceptions in a teachable agent framework. As far as Quinn's facial expressions, attribution theory postulates that primary emotions for outcomes may be refined according to the attribution a student makes for the outcome's cause (e.g., pride if the individual caused the outcome vs. gratitude if a teacher did) (Hareli & Weiner, 2002). Since attribution-related emotions may be only subtly different, Quinn expresses a single primary facial emotion (see Fig. 1, right), shown for about 15 seconds. Quinn can also highlight the attribution-specific emotion in the spoken message (e.g., "*That was right. Oh man, I am smart at math.* **I feel proud").**

Table 1: Sample 6 of the 12 Quinn attribution messages (suffix specifying emotion not shown)

Outcome	*Cause (Agent/Source)*	*Quinn's Message*
correct	I / ability	That was right. Oh man, I am smart at math.
correct	you / ability	Yay! I got that right because you are a good teacher.
correct	we / ability	That was correct! My gosh, We are good at this.
incorrect	I / effort	Oh boy. I got that wrong because I did not try hard to learn.
incorrect	you / effort	You did not put in much effort into teaching me that problem.
incorrect	we /effort	Dang it, that was wrong. We did not work hard to solve that problem.

Students' Perceptions of Quinn: User Study

In order to obtain data on students' interactions with the social robot Quinn, and in particular the different attributions it makes, we conducted a user study. Our specific research questions included:

(1) What are students' reactions to Quinn?
(2) How do students respond to Quinn's causal attributions for failure and success?
(3) Are some attributions more appropriate for fostering social interactions in a teachable agent framework?

To answer these questions we had students solve problems in TAG, because we wanted to afford students the opportunity to interact with Quinn and thus be able to ground their perceptions in their experience. We then used semi-structured exit interviews as our primary source of data. We chose this methodology as it had the potential to provide richer data on students' experiences than, for instance, affective surveys.

Materials

The study involved the following materials: two cheat sheets (domain and a system commands), an attribution questionnaire, the TAG problems, and a set of solution cards. The domain cheat sheet reviewed the target concepts related to plotting and translation; the system cheat sheet described the set of TAG commands. The attribution questionnaire probed student attributions through a series of multiple-choice questions. Eight of the questions proposed a hypothetical situation and asked students to select the choice that best fit their reaction

(from six choices representing the *effort/ability* and the *I/you/we* dimensions). Two questions included teaching-centric scenarios (e.g., *A friend that you have been tutoring in math has aced the math test*) and six were student-centric scenarios (e.g., *You have just received an A on your math test at school*). The TAG problems corresponded to two types: *plotting* of a point and *translation* of a point. The solution cards were 8x11 sheets of paper, one for each of the TAG problems; a given sheet was labeled with a TAG problem number on the front and a detailed description of the steps needed to generate that problem's solution in TAG on the back.

Participants

The study participants were 19 5th and 6th grade students (8 female) from a middle school in a large southwestern city. Students participated on site at their school, outside of regular classes but during regularly-scheduled classes (i.e., students individually left class to participate), and received $20. We chose students from grade 5 and 6 because these students already had some exposure to our target domain, but were not expert in it, as we identified by checking state standards and confirmed with pilot evaluations and discussions with teachers.

Procedure

Subjects (1) signed an assent form (~5 min.); (2) read the domain cheat sheet (~5 min.); (3) filled in the questionnaires (~15 min.); (4) were trained on how to use TAG (~20 min.); (5) used TAG to teach Quinn (45 min.); and (6) participated in a semi-structured interview (~20 min.). Sessions were conducted individually and were videotaped; two experimenters were present during each session. For the training phase, an experimenter followed a predefined script to go over TAG functionalities with each student. To implement the teaching framing, students were asked to *"tutor Quinn about how to solve geometry problems. The goal is for Quinn to learn enough so that it can solve all kinds of geometry problems. So when you are telling it how to solve a problem, think about what would be most useful for Quinn to know"*. Students were also told that they could refer to the cheat sheets and the solution cards and that it was up to them as to how they used these. Since prior work indicated that when peers are friends certain social behaviors are associated with learning (Ogan et al., 2012b), students were asked to pretend that Quinn was a long-time friend; following the teachable agent paradigm, they were also told that it is Quinn who gets the answer correct or not. Students then "taught" Quinn by working through geometry problems. Once a problem was solved, TAG provided feedback for correctness, and Quinn responded by attributing its success or failure to one of the target attribution dimensions. Since we are exploring students' reactions, a given attribution was chosen at random. For a given student, an attribution was never shown twice before all attributions were used. We manipulated whether students also heard Quinn express the attribution-specific emotion in its message (n = 9), to obtain data on students' reactions to receiving this information verbally and explicitly rather than by interpreting Quinn's emotion from its face. As the final step, students participated in a semi-structured interview led by the lead investigator. The interview questions probed students' reactions (e.g., on whether they felt they were teaching), but also focused on Quinn's attributions (both ones they heard in the teaching phase and any remaining ones they did not). To increase realism, during the interview, the experimenter played Quinn's messages for students using Quinn's voice. After the interview, students were debriefed, by being informed that Quinn's interventions were chosen at random and were not directed at them personally.

Results

The interviews were transcribed and the data was analyzed through qualitative description. Specifically, we iteratively derived codes from the data, organized these according to emergent themes, and refined them as needed. Our goal was to provide a qualitative summary of students' perceptions. We also analyzed student responses to the attribution questionnaire, by creating an attribution profile for each student based on frequency counts for the "source cause" (*ability* vs. *effort*, collapsing across *I/you/we)* and the "agent cause" (*I* vs. *you* vs. *we*, collapsing across *ability/effort)*, for positive and negative outcomes.

While we focus our analysis on students' reactions to Quinn attributions, we begin by mentioning overall perceptions of Quinn, its facial expressions, and the teachable agent framing. When asked about Quinn, not a single student expressed dislike for the robot, despite some of its negative attributions. Instead, students had positive reactions, many specifically mentioning Quinn's social behaviors. For instance, students said that the thing they liked the most about TAG was "how Quinn showed his feelings" (s1), Quinn because it "was cute" (12), and how "the robot was talking and stuff to me" (s19); s11 echoed this by stating that "it's cool that he has face emotions". While s14 mentioned that he got frustrated "a little bit […] when Quinn … got it wrong and he was getting mad at me", when asked if a Quinn who did not speak or show faces would be preferable, he responded it would "be worse because then you wouldn't know what he would be feeling". S17 mentioned that Quinn's attributions made him feel "good" and s9 wanted to "take Quinn home" because it was "helpful". As far as Quinn's facial expressions, some students explicitly mentioned liking Quinn's faces (s4, s7, s12, s14, s16, s19). For instance, s16 stated that the faces "make it more fun, cause when you get it right he [Quinn] will be happy", adding that this made him happy. In contrast, some students mentioned no preference for Quinn having

facial expressions (s2, s13, s15, s18). Students also did not express a clear preference for having Quinn verbalize its emotions (recall that for some students, Quinn suffixed an emotion following its attribution).

As far as our teaching framing manipulation, the majority of students felt they were teaching Quinn (s1, s3, s5, s7-s11, s13, s16, s18, s19). Some students mentioned that Quinn's attribution messages made them feel like this (s3, s8, s11, s16), e.g., "when she was saying positive things like I'm a good teacher" (s11), and when "he said that I taught him how to do it" (s16). S5 said it was because "I was showing him where ... the coordinates were" and s19 mentioned "by listening and stuff**"**. Other students did not buy into the teachable framing (s2, s4, s12, s15, s17). S2 said this was because "she [Quinn] knew where to go already". The remaining students said that it was because they were "controlling" Quinn and that Quinn initiating actions would make them feel like they were teaching Quinn more. One student felt "the robot was teaching me" (s14). We now present the attribution results organized by incorrect and correct outcomes.

Student Perceptions of Quinn's Attributions for Incorrect Outcomes

For the *I/we* dimensions for incorrect outcomes, students recognized the utility of effort over ability because "she [Quinn] knows she should try harder and she might get it" (s11). Students also commented on the fact that the effort attributions made them feel like a teacher (e.g., "next time like try to learn while I'm teaching" s18). In contrast, for the *you* dimension we did not see clear differences between ability and effort - details are below. The interview data for effort attributions is aligned with the attribution questionnaire data, in that the majority of students selected effort for the student centric (n = 12) and teaching centric (n = 15) questions. As far as the "agent cause" dimension (*I/you/we*) in the questionnaire data, for the student-centric questions, students attributed to themselves (*I,* n = 14) or to *we* (n = 3); for the teaching-centric questions, students attributed to themselves (*I,* n = 11), to *we* (n = 5) or to *you* (n = 3) (minor variations in N are due to missing data).

***I* (Quinn) dimension / *ability* + *effort*.** Most students did not appreciate Quinn's attribution to its ability upon an incorrect outcome. S14 said it was not realistic "because he is not really dumb at math". Other students had been taught to not attribute failure to ability, and transferred this to Quinn (s4, s5, s7, s12, s13, s17, s18). S5 said that "you shouldn't call yourself dumb", while S7 said that it's "not nice saying that to himself"; S13 echoed this, i.e., "he's putting himself down". Some students learned these sentiments from their parents (e.g., "my mom told me not to say this", s4). S8 felt that Quinn shouldn't take all the blame, because "it is both our fault". Some (s11, s17) didn't like the message "because she [Quinn] is putting too much pressure on herself - she thinks that she can't do it but we all know she can do it" (s11). Students also mentioned feeling sad for Quinn when they heard the message (s12, s17). When asked how the attribution would make them feel in terms of teaching Quinn, s12 answered "Quinn shape up!", while s17 wanted to "make it feel better" by giving it "an easy problem". In contrast, s3 liked the message, saying that it would make her want to "teach him more"; s10 echoed this sentiment. S6 said the message was funny, empathizing with it more than the effort attribution because "you can try hard and still not get it".

In contrast to Quinn's attribution to ability, more students appreciated Quinn attributing the negative outcome to effort (s4, s7, s9, s14, s18, s19), for instance "because he's being honest that he was not really listening" (s4). S5, who got the message appended with Quinn's attribution-specific emotion (guilt), also liked the attribution, but did not think Quinn should feel guilty, so as to not "sound hopeless". In contrast, s1 mirrored Quinn's guilt, i.e., "I would also feel guilty because I am teaching him", and that this would make him want to teach Quinn more. Several students, however, did not buy into the message (s1, s8, s13, s16). S1 did not believe that Quinn failed due to lack of effort, because "I'm doing everything for him to get him the right answer"; likewise s16 said it was "who controlled it didn't try hard [instead of Quinn]". Along a similar vein, s8 stated that "it wasn't her fault, it was mine", but also added that Quinn "could try a little bit harder".

You* (student) dimension / *ability* + *effort. When Quinn attributed the negative outcome to its teacher (i.e., the student), we did not see a clear difference in student perceptions between ability and effort and so collapse them here. Some students had a negative reaction (s1, s6, s9, s10, s14). S9 said he did not think Quinn was teasing him and that the message was "hurtful … [Quinn] needs to get … respectful"; s2 felt "sad that like I'm not teaching Quinn hard enough to get the problem right", but he also expressed that it "would make me try teach the robot more". Some students expressed frustration: S14 said the message made him "mad at him cause I'm trying to help him" and so might feel less like teaching Quinn, while s1 mentioned it hurt his "feelings a little" and as a result he "felt a little frustrated with [Quinn]", because Quinn "didn't say I'm sorry for anything"; likewise, s10 stated "I wouldn't help [Quinn]". S6 also said the attribution made him "a little mad", because "he is just blaming me". S11 had mixed feelings: Although she expressed preference for other attributions, she said this one "makes me feel like she doesn't think that I can do better but I think I can show her that I can", adding that it would both make her "a little mad" and also make her try harder.

Other students had more positive reactions (s2, s4, s8, s16-s19). S4 wanted "to help him more so that he would succeed"; this was echoed by s17. Some students said the message was fine because it came from a robot: S8 responded with "it was just a robot", and found the message amusing. Likewise, s16 said "it's a small robot, so I don't know how it could hurt your feelings", also mentioning the message was fair game if "you did

badly". S2 said that if it was coming from an "actual person … then I'll probably get mad but it's coming from a robot that doesn't really mean it". S17 agreed with Quinn that he didn't teach well because "I couldn't do translation" and while its message did not make him feel "bad", it did make him feel "guilty".

***We* dimension / *ability + effort*.** When Quinn attributed failure to itself and the student (*we*), some students mentioned liking the "we aspect" without responding specifically to agent causing it (s9, s11, s16, s12, s19). S12 said "it's okay with me because he said *we*, so it's not only him [and] it's not only me that did all the work". Likewise, s11 felt the message was appropriate because it "partly my fault and its partly her fault" qualifying with "I don't know if its all of her fault because she is new [..] and she hasn't done this a lot". S16 said the message made sense as Quinn "said *we,* because he followed me and I did it wrong, so me and his fault". In contrast, s3 said that Quinn was "kind of blaming" her for the outcome.

Reactions to the ability attribution in the *we* dimension were mixed. S15 felt it meant "that I have to be better at math" and make him try harder; when asked why him and not Quinn, he responded with "because I'm making it". S2 thought the message was "fine" but cautioned that "other people might not". In contrast, s11 did not like the attribution because she knew "she was fine at math". S17 felt the message unfairly blames both parties: "if you're saying you're bad, it's not like the one who's working with you is bad too". Although s18 found the attribution "funny", he also thought it felt "weird". Thus, in general, the majority of students explicitly mentioned preferring the effort attribution (s1, s2, s4, s5, s7, s9-s14). S9 empathized with the effort aspect, by saying that "sometimes at tests cause I'm so nervous I just guess". Others said this was because the ability one made them feel "something bad" (s4) or was "hurtful" (s12). S10, s13 and s14 elaborated on this sentiment by saying the effort attribution "wouldn't let the other person down […] and could help each other work harder" (s10); "well he's not putting anyone down he's just saying we did not try very hard on that" (s13) and "it's not like you suck at math […] you're just not working hard enough" (s14). S7 added that it made him feel like "if we try it again together we could figure it out". S4 said that the effort attribution would make him "try harder" over the ability one. S1 thought it was more realistic to attribute to effort "because well we've done a few problems, but it's not like we're really bad at it", a sentiment echoed by s5, i.e., "everyone is ok when working on math - no one is perfect". In contrast, s17 commented on the fact that Quinn did not know how hard she worked ("it's not like he can tell by looking") and so felt the effort attribution was "a little mean spirited".

Student Perceptions of Quinn's Attributions for Correct Outcomes

In contrast to incorrect outcomes, for correct outcomes the majority of students did not specify a preference for ability over effort for the *I* and *you* dimensions, but did prefer effort for *we* dimension. While students responded positively to all three of the *I/you/we* dimensions, they were especially enthusiastic about *we*. The student profiles extracted from the attribution questionnaire indicate that the vast majority of students selected effort for the student-centric (n = 18) and teaching-centric (n = 17) questions. As far as the "agent cause" dimension (*I/you/we*) in the questionnaire data, for the student-centric questions, students attributed most often to themselves (*I,* n = 12), followed by *you (*n = 2) and *we (*n = 1); for the teaching-centric questions, attributions were similar for *I* and *we* (n = 7 and n = 8) and less common for *you* (n = 3).

***I* dimension (Quinn) / *ability + effort*.** Many students responded positively when Quinn attributed success to either its ability or effort (s1, s4, s5, s7, s11, s12, s14, s16, s18, s19). For instance, s1 said Quinn "felt proud of me making him make the right choice" and s11 agreed with the attribution because "she [Quinn] is smart at math". Reactions to students who heard Quinn expressing the emotion (pride) in the message were mixed. S16 mirrored the pride ("I taught him and made him good at it") while s5 felt Quinn was "boasting a little"; this was echoed by s17 who said Quinn was "a little too proud". S12 who got the attribution without the emotion also felt like Quinn was bragging, adding "think of the other students". Thus, some students preferred the effort attribution (s4, s7, s11, s12), recognizing the utility of effort over ability, e.g., "if you try hard you usually succeed" (s4). S7 empathized with the effort attribution "cause that's like me, I work so hard until I finally get it", while s11 felt the attribution made her feel like she was teaching Quinn "because she tried harder to do that problem ... and got it right the second time". Several students could not pick between effort and ability attributions (s1, s5, s18), with s1 saying that "both have good feelings not against others". Students who did not like the attribution said it was because Quinn was taking all the credit. S14 responded by saying "*we* are smart at math"; s8 explained that "it makes me feel a little sad cause she's saying that she's the only one that did it". Likewise, s9 felt that Quinn was mean because "he didn't put me in that sentence".

***You* dimension (the student) / *ability + effort*.** When Quinn attributed its success to the teacher (i.e., student), the majority of students did not specify a preference between the effort and ability, and most reacted positively to both. To illustrate, students said "that is really, really – it's nice and kind of him to say" (s1), "that's very nice" (s12), "that's a good one" (s14), "I feel really good about that" (s4), that it felt "like I've accomplished something" (s13). Students also described how the attribution made them want to teach Quinn. S4 said "because she [Quinn] said I was a good teacher and I didn't want to say like - no I don't want to do this anymore". In response to the effort attribution, s1 said that he liked the message because "I taught him … it's like we did it together." In contrast, s13 wanted Quinn to acknowledge that they were working hard and so

preferred the effort attribution, while s3 did not like the effort attribution as it meant Quinn was not putting in effort itself, i.e., "that would say that he just listen to me for the answer and he didn't do anything".

***We* dimension / ability + effort.** S14 liked attributions for success that included them and Quinn (i.e., *we*) because "if I was the only one who did it he'd feel bad … so when it's both of us it's like we're both getting the same share". S1 said *we* was best because "he was really happy that we both did it"; s6 "I like that out of all of them". S18 stated that "expresses a lot because the person controlling him feels great because he was teaching him and it makes them feel good to keep trying". S10 said in response "she [Quinn] was right that we are good at this, and that she tried harder at the question and she knew what she was doing". Likewise, students also liked the effort attribution (s1, s6, s7, s17, s19), because it made them "feel good" (s6). When asked to compare effort to ability attributions, the majority preferred effort (s3, s6, s9-s12, s17 - s18). S3 said this was because the effort attribution "means we learned… if we were to say we are good at this, then that means we already knew it." S18, s6 and s9 liked that the message acknowledged their effort (e.g., "geometry is hard and I had to help Quinn… so we both worked hard", s6), with s9 adding that because they worked hard Quinn "is happy". S10 picked the effort attribution, but also added that the ability one "gets my hopes up". S17 said that both attributions were cheerful, but also picked the effort one. Other students did not have a preference between the two attributions (s5, s8, s14), saying, for instance, that they both "sounded good".

Discussion and Future Work

Weiner's theory proposes that students should attribute to high ability or high effort for correct outcomes, and to low effort for incorrect outcomes (Försterling, 1985). The majority of our results for the type of ability and effort attributions students feel *Quinn* should make mirror this finding. In the future, we plan to use a computational model that assesses a student's effort and takes into account prior history of success and failure to select Quinn's attributions. For instance, the robotic agent could attribute success to effort using the *you* dimension ("you tried hard to teach me") when the student struggles though an especially challenging problem. We now explore how Quinn might take on these adaptive attributions.

For incorrect outcomes, our analysis suggests that attributions to effort in the *I* (Quinn) and *we* dimensions are safe for a social robot to express, in terms of maintaining positive student affect and motivation to teach. Prior to the study, we anticipated students may have preferred to shift the responsibility of an incorrect outcome on the agent, as found in non-educational settings (Groom et al., 2010). However, the pre-study questionnaire data indicated that for incorrect outcomes, for both student-centric and teaching-centric questions, students most often attributed failure (and success) to themselves. Interestingly, interview data revealed that when actually teaching, the *we* dimension got the most enthusiastic responses for incorrect outcomes – this finding is aligned with the fact that teaching is inherently a social activity. Students appreciated spreading the blame attribution between themselves and Quinn, which suggests students felt responsible for the robot, and so did not want it to take all the blame. When Quinn instead attributed failure to the student teaching it (the *you* dimension), most students responded with frustration, stating that the message was not motivating them to teach Quinn. Prior work with students in peer-to-peer interactions showed face threatening moves of this type were in some cases associated with learning (Ogan et al., 2012b), and in fact we did find that some students had a positive reaction, even saying that the message made them want to teach more. A promising avenue, proposed by one of the students, is to have Quinn exhibit playful affect when stating such messages (this student suggested Quinn should stick out its tongue). Thus, our results suggest that this *you* dimension should only be used for students who will interpret it as a playful challenge, rather than as unproductive rudeness. As far as the *I* dimension, students reacted negatively when Quinn attributed incorrect outcomes to its own low ability. Prior work suggests that this should elicit either sympathy or contempt in the observer (here, the student teaching Quinn). While one student did express sympathy for Quinn, for the most part students instead expressed irritation (and possibly contempt) that Quinn was putting itself down. When Quinn attributed failure to low effort, we did not see the reaction forecast by prior work, namely anger. Instead, most students reacted positively, some adding that the attribution would want to make them teach Quinn more. *In sum, these results suggest that for incorrect outcomes: (1) students prefer we attributions that share the blame between them and Quinn, but (2) you attributions might be motivating for students who see them as a playful challenge, and (3) I attributions could motivate students to try teach Quinn the relevant concepts*.

For correct outcomes, our data indicates a slightly different picture in terms of design recommendations for social teachable agents, in that all six attribution types across the ability/effort and I/you/we dimensions were positively received. Thus, having a robotic agent attribute to all dimensions, with varying degrees of frequency, could increase believability and student bonds with the agent. In particular, the *we* attribution was very popular, with students appreciating Quinn, recognizing their role in the successful outcome (a correct solution). However, the *you* and *I* dimensions also fared well. For the *you* dimension, students appreciated hearing compliments about their teaching, suggesting that they bought into the teaching framing to some extent. Likewise, both *I* attributions were well received, and compared to incorrect outcomes, for correct ones there was not as strong a negative response to Quinn's ability attributions, with some students mirroring Quinn's pride

expressed in this attribution. While the effort attributions for the *I* dimension may have been more positively received than ability ones, having Quinn attribute to both ability and effort could add to the realism of its design. It is interesting to note that students responded enthusiastically to the effort attributions made by Quinn upon a correct solution, for instance empathizing with Quinn because they also had to try hard to learn. These reactions are in opposition to prior work indicating that when students attribute success to effort, they may be seen as less capable (Försterling, 1985). *To summarize, upon a correct answer: (1) students perceive all of Quinn's attributions positively, but (2) you attributions might be particularly effective for students who need to boost confidence in their teaching abilities, and (3) I attributions to effort might encourage students to take on effort attributions for their own problem-solving.*

To conclude, prior work has shown that when students make appropriate attributions with respect to themselves, they persist more and learn better. There is much less work exploring the impact of others' attributions has on observers who overhear these. Here, we took the first steps in filling this gap by investigating students' perceptions of a teachable robot that attributes success and failure to different causes, showing that students' perception of such a robot is influenced by various attributions it makes.

References

Baylor, A., & Kim, S. (2008). The Effects of Agent Nonverbal Communication on Procedural and Attitudinal Learning Outcomes. In *Proc. of Intelligent Virtual Agents*, 208-214.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at, V. (2005). *Learning by teaching: a new agent paradigm for educational software. Applied Artificial Intelligence*, 19(3-4), 363-392.

Chase, C., Chin, D., Oppezzo, M., & Schwartz, D. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, 18(4), 334-352.

Chi, M. T. H. (2009). Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1, 73-105.

Försterling, F. (1985). Attributional retraining: A review. *Psychological Bulletin*, 98(3), 495-512.

Groom, V., Chen, J., Johnson, T., Kara, F. A., & Nass, C. (2010). Critic, compatriot, or chump? Responses to robot blame attribution. In *Proc. of Human-robot interaction*, 211-218.

Gulz, A., Haake, M., & Silvervarg, A. (2011). Extending a Teachable Agent with a Social Conversation Module - Effects on Student Experiences and Learning. In *Proc. of AIED 2011*, 106-114.

Hareli, S., & Weiner, B. (2002). Social Emotions and Personality Inferences: A Scaffold for a New Direction in the Study of Achievement Motivation. *Educational Psychologist*, 37(3), 183-193.

Kanda, T., Shimada, M., & Koizumi, S. (2012). Children learning with a social robot. In *Proc. of HRI*, 351-358.

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain System. *International Journal of Artificial Intelligence in Education*, 18(3), 181-208.

Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., & Paiva, A. (2010). "Why Can't We Be Friends?" A Game Companion for Long-Term Interaction. In *Proc. Of Intelligent Virtual Agents,* 315-321.

Muldner, K., Lozano, C., Girotto, V., Burleson, W., & Walker, E. (2013). *Designing a Tangible Learning Environment with a Teachable Agent.* In Proc. of Artificial Intelligence in Education , 299-308.

O'Malley, C., & Fraser, D. S. (2004). *Learning with Tangible Technologies*, Literature Review series, report 12.

Ogan, A., Finkelstein, S., Mayfield, E., D'Adamo, C., Matsuda, N., & Cassell, J. (2012). "Oh, dear Stacy!" Social interaction, elaboration, and learning with teachable agents. *In Proc. of CHI 2012*, 1381-1390.

Ogan, A., Finkelstein, S., Walker, E., Carlson, R., & Cassell, J. (2012). Rudeness and Rapport: Insults and Learning Gains in Peer Tutoring. *In Proc. of ITS 2012*, 11-21.

Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. *In Proc. of Human-robot interaction*.

Reif, F., & Scott, L. A. (1999). Teaching scientific thinking skills: Students and computers coaching each other. *American Journal of Physics*, 67(9), 819-831.

Roscoe, R. D., & Chi, M. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77(4), 534-574.

Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. *In Proc. of CHI'10*, 1613-1622.

Woolf, B., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., & Christopherson, R. (2010, June 2010). The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities. *In Proc. of ITS'10*, 327-337.

Acknowledgements

This research was funded by NSF 1249406: EAGER: A Teachable Robot for Mathematics Learning in Middle School Classrooms and by CAPES Foundation, Ministry of Education of Brazil- DF 70040-020.

Characterizing a New Dimension of Change in Attending and Responding to the Substance of Student Thinking

Jennifer Richards, Andrew Elby, Ayush Gupta, University of Maryland, 2311 Benjamin Building, College Park, MD 20742
Email: jrich@umd.edu, elby@umd.edu, ayush@umd.edu

Abstract: "Responsive teaching," in which teachers attend and respond to the substance of students' ideas, is central to facilitating student learning through engagement in authentic disciplinary practices. In characterizing teachers' progress toward greater responsiveness, researchers typically code teachers' attention as shifting toward the intellectual content (substance) of students' ideas and away from other foci such as students' correctness. These schemes, however, do not distinguish between different aspects of the substance of students' ideas. In this paper, we argue that a science teacher, Mr. S, demonstrates progress not by shifting toward greater attention to "substance," but rather by shifting in the facet of student thinking to which he primarily attends and responds. He shifts toward attending to causal stories (mechanistic explanations) and away from causal factors (potentially relevant variables). We argue that such shifts toward more sophisticated epistemic practices should be targets of professional development and of the assessment of responsive teaching.

Introduction

When teachers attend and respond to students' ideas and seek to draw out or connect them with important aspects of the discipline, students demonstrate enhanced conceptual understanding (e.g., Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Pierson, 2008) and experience rich opportunities to engage in disciplinary practices, such as explanation-building and argumentation in science (e.g., Berland & Reiser, 2009; Duschl & Gitomer, 1997). Ball (1993) has described this sort of teaching as involving "twin imperatives of responsiveness and responsibility" (p. 374) – focusing on and grounding instruction in students' ideas, while helping them learn important disciplinary ideas and practices. In science, Hammer and van Zee (2006) highlight the importance of teachers focusing on various beginnings of science in what students are saying and doing. Take the following example they discussed: A student says it gets hotter in the summer because the earth is closer to the sun. Although this idea is incorrect and widely considered to be a common student misconception about the seasons, Hammer and van Zee emphasize the scientific features of the explanation – its mechanistic nature, tangibility, and consistency with other information the student knew. These are a few examples of scientific aspects teachers could note and promote in students' reasoning; Hammer and van Zee describe numerous others (e.g., anticipation of counterarguments, clarity of expression, etc.).

Characterizations of favorable change in attending and responding to the substance of student thinking, however, primarily emphasize how closely teachers focus on students' meanings with little attention to the sorts of discipline-specific features they notice. Researchers tend to focus on the specificity with which teachers attend to students' ideas (e.g., Jacobs, Lamb, Philipp, & Schappelle, 2011; van Es, 2011), the stance teachers take toward students' ideas (e.g., Crespo, 2000; Goldsmith & Seago, 2011), and/or the types of follow-up moves teachers make in response to students' ideas (e.g., Brodie, 2011; Pierson, 2008). These foci foreground teachers' treatment of students' ideas – in some cases, in the context of specific disciplinary domains, like elementary school students' mathematical problem-solving strategies – but do not clearly address the range of disciplinary aspects teachers may attend to in what they hear.

In this paper, our primary aim is to bring discipline-specific considerations into the discussion of change in attending and responding to the substance of student thinking. Drawing on two similar lessons taught by middle-school science teacher "Mr. S" in successive years, we argue that part of what constitutes the favorable change in Mr. S's attention and responsiveness to student thinking between April 2010 and March 2011 is the aspects of scientific reasoning and explanation Mr. S foregrounds with respect to students' ideas in each case.

Literature Review: Identified Dimensions of Favorable Change in Attending and Responding to Student Thinking

We begin by reviewing the dimensions along which researchers describe favorable change in attending and responding to student thinking, demonstrating that these dimensions are largely free of discipline-specific features.

A common consideration in evaluating responsiveness is whether teachers' descriptions of students' reasoning are a) general and draw on superficial aspects of the ideas or b) specific and draw on details and nuances within the ideas, with the latter considered more responsive (e.g., Crespo, 2000; Fennema et al., 1996;

 © ISLS

Franke, Carpenter, Fennema, Ansell, & Behrend, 1998; Jacobs et al., 2011; Kazemi & Franke, 2004; Levin & Richards, 2011; van Es, 2011). For instance, in a teacher work group in which teachers were expected to share how their students approached a particular mathematics problem, Kazemi and Franke noted that early on, teachers focused on whether students' strategies were correct or not but paid little attention to the specifics of students' solutions. Similarly, Crespo noted that the preservice teachers in her study initially made claims about student understanding that were not grounded in much evidence, but later, teachers' "comments revealed greater attention towards the meaning of student's mathematical thinking rather than surface features" (p. 170). Similarly, in van Es's framework, part of what distinguishes Level 1 (baseline) from Levels 3 (focused) and 4 (extended) is that teachers in Level 1 "form general impressions of what occurred" and "provide little or no evidence to support analysis," whereas teachers in Level 3 or 4 "refer to specific events and interactions as evidence" (p. 139).

Another way in which researchers characterize favorable shifts in attending and responding to student thinking is in movement from an evaluative stance (looking only for correctness) to an interpretive stance (making sense of students' ideas) (e.g., Brodie, 2011; Crespo, 2000; Empson & Jacobs, 2008; Goldsmith & Seago, 2011; Levin & Richards, 2011; Pierson, 2008; van Es, 2011). In teacher work groups, researchers analyze the types of comments teachers make when looking at student artifacts. For instance, Goldsmith and Seago described how early mathematics teacher work groups interpreted student work in light of the correct answer, whereas later they looked for the logic in students' own solutions. Revisiting van Es's framework, stance is also integrated into the different levels. For example, in Level 2 (mixed), teachers "provide primarily evaluative with some interpretive comments" (van Es, p. 139), whereas in Levels 3 and 4 teachers shift to entirely interpretive comments.

In situ in the classroom, researchers interpret the teacher's stance based on the follow-up moves the teacher makes in response to students' ideas. For example, Empson and Jacobs (2008) define a progression in listening expertise that moves from directive listening (where the teacher focuses on alignment between a student's idea and an expected response and actively seeks to elicit the expected response) to observational listening (where the teacher passively listens to students' ideas without seeking to extend them) to responsive listening (where the teacher actively probes students' ideas and seeks to understand and build on the details). Similarly, Brodie (2011) and Pierson (2008) differentiate between follow-up moves that are corrective versus interpretive in nature. In these examples, the shift is in how teachers view and thus work with students' ideas and involves a change from seeking to evaluate students' ideas to seeking to understand them in more depth.

In summary, the shifts in attention/responsiveness described above depict movement from a) evaluating students' ideas, focusing on surface features to determine alignment with expected responses and making follow-up moves to push students in particular directions, to b) interpreting students' meaning, focusing on the details of students' ideas and making follow-up moves to elicit more information from students. While these are important aspects of responsive teaching, they do not shed much light on how teachers are attending and responding to specific disciplinary aspects of students' ideas. In this paper, we illustrate and tease apart two particular ways in which teachers' interpretive follow-ups to students' ideas can intersect with authentic disciplinary reasoning.

Data Sources and Analytical Approach

The data in this paper come from a professional development project aimed at helping fourth through eighth grade teachers promote inquiry teaching and learning in their science classrooms. Teachers voluntarily apply and may continue in the project for multiple years. As part of the project, teachers attend a two-week summer workshop in which they engage in their own minimally-guided inquiry, watch classroom video of students discussing scientific phenomena, and collaborate on other issues related to inquiry teaching and learning in the classroom (i.e., assessment, lesson planning, etc.). During the school year, teachers work one-on-one with members of our research team to facilitate scientific inquiry in their classrooms and attend bimonthly small group meetings with other teachers and members of the research team.

Our research team identified Mr. S – currently in his fourth year of participation in the project – as someone who came to consistently facilitate rich scientific discussions in his classroom, many of which we have videotaped. The two selected episodes in this paper come from Mr. S's seventh-grade classes at a Title I middle school in which 65% of the students identify as Hispanic, 30% as African American, and about 35% are classified as having limited English proficiency (1).

Specific features of this pair of episodes made them an ideal naturalistic setting for thinking about different scientific aspects teachers may attend to in students' ideas. In many respects, the episodes are similar – they feature the same teacher teaching the "same" lesson in consecutive years (April 2010 and March 2011). In both episodes, Mr. S posed the same basic question: If you're walking with keys, and you want to drop the keys into a container sitting on the floor, should you release the keys before the container, over the container, or after the container? Students posed sensible reasons for each option, and Mr. S entertained a range of possible answers. Yet what Mr. S foregrounded in students' explanations in each episode differed. Our research team had

previously noted that Mr. S's own explanations of scientific phenomena during the summer workshops varied in nature, at times identifying the causal factors responsible for the phenomena and at other times fleshing out more mechanistic explanations for how phenomena occurred. We noted that these different explanatory approaches seemed evident in his facilitation of this pair of episodes.

Our first analytical step was to fully transcribe the two videotaped episodes, each approximately fifteen minutes in length. The transcript captures pauses and emphases in participants' speech, drawing on transcriptional notations from Sacks, Schegloff, and Jefferson (1974): pauses in speech are indicated by long dashes (representing a beat) or (pause) (indicating a longer pause). Moments when a participant cuts himself off are represented by short dashes, and moments when a participant extends a word are represented by repeated colons in the middle of the word (e.g., "thi:::nk"). Emphases in speech are indicated by either underlined or capitalized words, with the latter representing increased volume specifically. Combinations of emphases and colons reflect pitch change in the course of a word, () indicates that the speech could not be deciphered, and actions are described inside double parentheses (( )).

We then compared Mr. S's attention and responsiveness to students' ideas in the two episodes in the following manner. We focused on exchanges in which common ideas came up in both episodes or in which Mr. S followed up with students extensively, because these sorts of exchanges were likely to provide useful points of comparison. We drew on three kinds of evidence to unpack what Mr. S was foregrounding during these exchanges:

- How Mr. S revoiced students' ideas (O'Connor & Michaels, 1993) – emphases in his summaries suggested what he primarily attended to (e.g., "Maybe GRA::vity. GRA::vity" [April 2010] vs. "Gravity's pulling it down" [March 2011])
- How and when Mr. S pressed on students' ideas (Brodie, 2011) – questions Mr. S asked students indicated what he wanted them to flesh out (e.g., "So you're saying some kind of forward motion based on what?" [April 2010] vs. "Why will the keys go fast too?" [March 2011])
- When Mr. S made verbal and nonverbal bids to close the conversation (Schegloff & Sacks, 1999; Stivers & Sidnell, 2005) – accepting students' ideas as sufficiently articulated demonstrated what he found satisfactory (e.g., moving to another idea after a student identified wind as influential vs. after a student explained *why* wind was influential)

Evidence from other data sources (e.g., debrief conversations with Mr. S, recollections from small group meetings, stimulated recall/reflection interviews) triangulated with our interpretation of what Mr. S was foregrounding in each episode. Due to space constraints, we only review evidence from the episodes themselves in this paper, but more information (including full episode transcripts) can be found in Richards' (2013) dissertation work.

Findings

Our analyses demonstrate that Mr. S foregrounded different aspects of students' scientific reasoning in his attention and responses in April 2010 versus March 2011. In the first episode, Mr. S foregrounded students identifying *causal factors* responsible for the motion they predicted. In the second episode, Mr. S foregrounded students articulating *causal stories* for the motion they predicted, fleshing out how and why the object would move the way it did. This shift in attention from causal factors to causal stories represents a favorable change in the sophistication of explanation Mr. S attended to and pressed students for in the context of the key drop question.

Mr. S Foregrounded Causal Factors in the First Episode

In the first key drop episode in April of 2010, Mr. S primarily attended and responded to a particular form of scientific knowledge in students' ideas – their identification of the causal factors or force-like entities responsible for the motion they predicted. In general, if the factor causing the motion was not apparent in a student's explanation, he pressed the student to articulate it; if the factor was apparent, he accepted the student's response. Here, we provide two in-depth examples to illustrate Mr. S's focus on causal factors and cite supporting evidence from other exchanges throughout the episode.

The following exchange occurred well into the discussion and was one of the longest continuous exchanges Mr. S had with an individual student during the episode. The student, Suri (all names are pseudonyms), provided his sense of when it would be best to drop the keys, if you're running fast:

1. Mr. S: Okay, Suri, you want to respond to that or add something to the discussion?
2. Suri: Yeah, I'm like, if you're running, you feel like the wind is pushing you back.
3. Mr. S: So you're saying as you're going fast, faster, you're also feeling some pressure, some air, pushing back against you.
4. Suri: So my drop, um, is from above or after.
5. Mr. S: Above or after because of what?

6. Suri: Because if the wind is working in a different direction than you, you're running and () ((moves one hand forward and the other in the opposite direction on top)).
7. Mr. S: So when you, when you're saying, when you're running fast, there's some pressure coming up against you, coming against you?
8. Suri: Mm-hmm.
9. Mr. S: What is that? (pause) What do you think that is? (pause) So you're saying there's a pressure, there's something pushing back against you. ((faces board, writes)) There's a push back. And, so that push back, when you release the keys, what is it going to do to the keys?
10. Suri: They're gonna drop backward.
11. Mr. S: They're going to drop back. Okay, okay. Um, now, what are some-

Throughout the exchange, Mr. S attended and responded to Suri's idea – he maintained his focus on Suri's idea and pressed Suri to say more. However, there are nuances in the ways Mr. S interacted with Suri that highlight Mr. S's emphasis on causal factors. For instance, after Suri provided his initial explanation and indicated that he would drop the keys above or after, Mr. S asked Suri, "Above or after because of what?" (line 5). The fact that Mr. S had already revoiced Suri's explanation in line 3 and the wording of the question in line 5 suggest that Mr. S may have been looking for Suri to further specify the particular factor he thought was in play. Instead, Suri reiterated his story of the wind "working in a different direction than you" (line 6), and Mr. S again acknowledged Suri's story but pressed for the responsible factor: "What is that?... What do you think that is?" (line 9). Note here that Mr. S attended to the causal story Suri provided about the wind working in a different direction and pushing back against you – this aspect was not completely absent. Yet what Mr. S pressed for was Suri's identification of the causal factor involved.

At several other times throughout the episode, Mr. S also pressed for or attempted to elicit specific factors or forces underlying the motion students described. For example, early in the discussion, one student, Jack, talked about the keys falling straight down because of their weight. Mr. S responded in part by asking, "What force will cause it to go straight down?" and excitedly accepted the response of gravity ("Maybe GRA::vity. GRA::vity"). Another example occurred when Katherine talked about the keys going backward if you're going fast. In response, Mr. S asked, "If I'm going fast, why would that cause the keys to go backwards? What, what force, what would cause the keys to go back?" His reframing of the question from *why* the keys would go backward to *what force* would cause the keys to go backward, and his subsequent summary that Katherine "said something about the wind," reflected his emphasis on causal factors.

Further evidence of Mr. S foregrounding causal factors in students' ideas comes from a close look at another exchange around an idea that came up in both key drop episodes – that the speed of the runner would make the keys move forward. In the first key drop episode, a student, Diane, related this scenario to what would happen if you were to jump out of a racecar:

12. Mr. S: Why before, Diane?
13. Diane: Because I thi:::nk that – well, let me try to give you an example, li:::::ke ((loudspeaker interruption)) I think, like, when you're racing? Like, you're in a racecar? And then, you know, let's say you have to () on fire or something? So when you're trying to land on the grass – because you're not going to get there right when you're at the grass or else you're gonna- because the car's fast, and you're going fast too. You gonna, like, get on the mud or something, so you're going to have to go before, so you know, you could, you know what I mean?
14. Mr. S: So what do you mean is that there's some kind of forward motion?
15. Diane: Yeah.
16. Mr. S: ((faces board, writes)) Okay. So you're saying some kind of forward motion based on what?
17. Diane: On the speed of the person who ().
18. Mr. S: So based on sp::eed, right?

Again, Mr. S attended and responded to what Diane was saying. In line 12, Mr. S's question ("Why before?") elicited a detailed causal story from Diane. His follow-ups, however, did not acknowledge Diane's specific example, but rather clarified the kind of motion she implied (line 14) and pressed Diane to identify the causal factor responsible for the motion (line 16). His verbal emphasis on Diane's identification of "sp::eed" as the relevant causal factor (line 18), followed by his moving on to another student, suggests this is the kind of explanation he was looking for.

Mr. S Foregrounded Causal Stories in the Second Episode

When Mr. S explored the same question with another group of students in March of 2011 during his second year in the project, he attended and responded to a different form of scientific knowledge in students' ideas – their causal stories of what they thought would happen. This foregrounding involved his continued pursuit of different stories and more mechanistic detail from students. We provide two illustrative examples.

As the discussion started, many students thought you should drop the keys over the container in order to get them in. Yet they offered multiple kinds of explanations, including restatements of their conclusions and problematic alternatives (e.g., "Because if we drop it before or after the container, it won't get in the container") and appeals to the skill of the person dropping the keys (e.g., "Some people have bad aim, so they can't even aim towards the trash can"). Among these explanations was the following causal story from a student, Cooper:

19. Mr. S: Um, Cooper?
20. Cooper: Um, above?
21. Mr. S: Above.
22. Cooper: Because like the gravity, like, when you put it up, it goes down.
23. Drake: It's heavy.
24. Mr. S: Cooper said that because it's heavy, what happens, Cooper, I have to, I have to drop it-
25. Cooper: No, gravity puts, like, pulls it down.
26. Mr. S: So, because gravity's pulling it down.

Here, Cooper offered both a causal factor and how it works – gravity as the factor that works by pulling things down (line 25). Even though Mr. S momentarily conflated Cooper's idea about gravity with Drake's idea about heaviness, note the follow-up question Mr. S asked – "what happens?" (line 24). Mr. S was not satisfied with the identification of the relevant causal factor; rather, he asked Cooper for more of a narrative and, after Cooper's response, verbally emphasized what gravity does (line 26, "gravity's pulling it down").

As students continued to offer different kinds of explanations, Mr. S returned to Cooper's explanation, emphasizing its causal nature: "So now let's, we want to get back to – why, why above? Cooper, you had some explanation why, what's the reason for it?" Mr. S recapped Cooper's response for a third time as he asked students for other reasons why you should drop the keys over the container: "Are there any other reasons why I should drop it above the container, other than Cooper said, the gravity's gonna pull it down. Why else might I drop it above the container?" This repeated emphasis on "why" and the causal story (what gravity does) instead of just the causal factor (gravity) suggest that Mr. S was looking for students to provide causal story-like explanations.

This push beyond causal factors became most apparent in an exchange with Chavez about the speed of the runner making the keys move forward:

27. Chavez: If you do it before, it'll go directly in? But if you do it like, like-
28. Mr. S: Why do we have to do it before again?
29. Chavez: Because it'll go, like, IN, like the keys will go in the trash can or the thing will go in the trash can.
30. Mr. S: What will cause it to go in the trash can if we drop it before as opposed to over, because earlier you said over?
31. Chavez: Like, like, like, like, like the speed of the keys also I guess coming off.
32. Mr. S: The speed of the- so the keys have speed?
33. Chavez: Because you're walking, no, because like you're walking? (pause) And like, and like since you're walking fast, like, I guess the keys will also go fast too?
34. Mr. S: The keys will go fast too?
35. Chavez: (pause) Yeah.
36. Mr. S: Why will the keys go fast too?
37. Chavez: I don't know!
38. Mr. S: I released the keys, wouldn't the keys just be there?

Recall how the exchange between Mr. S and Diane went the previous year when the idea of speed came up. Mr. S emphasized Diane's identification of speed and moved on to another student. Here, there are notable differences in Mr. S's response, despite the parallels between Diane's idea that "the car's fast, and you're going fast too" [April 2010] and Chavez's idea that "since you're walking fast… the keys will also go fast too" [March 2011]. First, Mr. S did not simply accept the idea of speed; he started to repeat it (line 32) but then reflected the idea back to Chavez with a questioning intonation (lines 32, 34). Second, Mr. S pushed Chavez to fill out an additional part of the story by asking, "Why will the keys go fast too?" (line 36). This question, followed by Mr. S's counterpoint that the keys might "just be there" once they're released (line 38), indicates that Mr. S was interested in more than the identification of speed as a causal factor. He was also interested in Chavez fleshing out a causal story for how the keys would still have speed after they'd been released.

Thus, although Mr. S attended to both causal factors and causal stories to some extent in both episodes, we can see that he foregrounded one or the other in each case. We now turn to a discussion of why Mr. S's foregrounding of causal stories represents a favorable change over his foregrounding of causal factors.

Discussion: Considerations of Explanatory Sophistication in Science

Work in science education (e.g., Chinn & Malhotra, 2002; Russ, Scherr, Hammer, & Mikeska, 2008; Sandoval, 2003; Windschitl, Thompson, Braaten, & Stroupe, 2012) emphasizes the importance of students constructing causal explanations for phenomena. For instance, Chinn and Malhotra drew on work from the psychology, sociology, philosophy, and history of science to argue that one aspect of authentic inquiry is "the development of theoretical mechanisms with entities that are not directly observable" (p. 186). Sandoval's analysis of causal coherence in students' scientific explanations also focused on causal mechanisms, how students chain causes and effects to create coherent explanations. In creating and developing causal stories of how or why something happened, students engage in a practice that is arguably at the core of science.

When possible, fleshing out causal stories is a more sophisticated form of scientific explanation than simply identifying relevant causal entities or factors (Russ et al., 2008; Windschitl et al., 2012) (2). For instance, Russ et al. developed a framework for analyzing students' mechanistic reasoning, adapted from philosophy of science studies on the work of scientists. In Russ et al.'s framework for analyzing mechanistic reasoning, identifying entities and properties and actions of entities relevant to the target phenomenon is one component, but more sophisticated mechanistic reasoning involves creating a coherent explanation of how these activities or properties bring about the target phenomenon. In other words, identifying causal factors contributes to but is less sophisticated than telling causal stories, which requires consideration of how the factors behave and interact with each other over time.

Moreover, in the key drop conversations, students demonstrated the ability to engage in causal storytelling that could have been capitalized on both years. Consider Diane's racecar example from the earlier episode. Although she offered details that could have been drawn out further, Mr. S summarized her idea as having to do with speed and forward motion. Although this summary was coherent with Diane's idea, it quickly slotted her idea as a certain kind of thing rather than permitting further exploration. Diane's idea might have played out differently in the second episode. Judging from the Chavez exchange, Diane might have been asked to explain why you would still be going fast once you jumped out of the car. Rather than assuming that the same mechanisms were in play in the key drop scenario and the racecar example, the relationship between the situations might have been interrogated. In short, various aspects of Diane's explanation might have received deeper attention, and productively so for students' learning through engagement in scientific inquiry.

Thus, the favorable change seen in Mr. S's attention and responsiveness can be characterized *not* simply as greater attention to the substance of students' reasoning, but rather, by which aspects of the substance of student reasoning he foregrounded in his attention and responses. This shift in foregrounding from *causal factors* to *causal stories* does align with a dimension of favorable change noted above, namely the specificity with which Mr. S attended to students' ideas. Foregrounding causal stories necessitates attention to details of students' explanations in a way not required by attention to causal factors. However, with respect to the stance Mr. S took toward the ideas he heard, as reflected in his follow-up moves, we do not see differences between the episodes. In neither case did Mr. S direct the conversation toward the correct answer, nor did he listen passively; he was engaged in interpreting students' ideas in both episodes. Similarly, drawing on Brodie's (2011) scheme, the most frequent types of follow-up moves in both episodes were the reform-type moves of maintaining focus on students' ideas and pressing for more information. What *was* distinct between the episodes was the kind of information Mr. S pressed *for*.

Conclusion and Implications

We demonstrated that the shift seen in Mr. S's attention and responsiveness to the substance of student thinking between the key drop episodes hinged on which aspects of scientific explanation he foregrounded in relation to students' ideas. In his first classroom implementation of the key drop inquiry, Mr. S foregrounded students' identification of the causal factors or force-like entities responsible for the motion they predicted, such as *gravity* moving the keys down, or *speed* resulting in the keys' forward motion. Mr. S's subsequent classroom implementation of the key drop inquiry the following year, though, involved a more sophisticated foregrounding – students' articulation of causal stories of what they thought would happen. Here, mechanism was more of an emphasis, e.g., gravity pulling the keys down, or a lingering question about how the keys still have speed once they're released.

We recognize the limitations inherent in a case analysis of a single teacher, and as such, we do not claim that the specific shift seen here extends beyond this case, nor that discipline-specific considerations are *always* relevant in characterizing teachers' attention and responsiveness to student thinking. Rather, we see this case as a proof of concept that in some cases, as in the case of Mr. S, there is a disciplinary dimension to responsive teaching that is not captured by current characterizations in the literature. Additionally, space constraints prevent us from exploring two related issues: the reasons behind Mr. S's shift, and evidence that the shift represented stable change in practice rather than a mere "fluctuation." In other work we explore these issues (Richards, 2013). However, the data and analysis we presented here are sufficient to support our main argument that nuanced considerations of disciplinary authenticity and productivity, when evident in the data,

should inform notions of what counts as progress in responsive teaching—specifically, that shifts in attention to more sophisticated aspects of the substance of student thinking should count as progress toward greater responsiveness.

To conclude, we consider the implications of this work for professional development and research in which students' ideas are at the core of teachers' attention.

Implications for Professional Development

When students' ideas are central to professional development efforts, a critical topic for ongoing discussion should be the various disciplinary aspects that participants (including professional developers) note with respect to student thinking. After spending time making sense of students' ideas, participants could be asked to reflect on what is scientific about what students are saying and doing. For instance, in this case, Mr. S's sense that mechanism came up more in the second summer workshop likely influenced what he paid attention to in students' explanations, and how and when he pressed students to fill in gaps. More explicit discussion of such disciplinary aspects could help teachers open up space for students to explore and develop a deeper sense of a given discipline and what it means to engage and participate in ways that disciplinary experts do, in a sense becoming local disciplinary experts themselves.

Moreover, more explicit discussion would promote metaawareness in teachers of what they are foregrounding in given moments. Throughout his participation in the project, Mr. S has tended to use the terms causal factors and causal stories interchangeably, suggesting that he may not have been aware of his different foregroundings in the episodes. Thus, it is important to note how pervasive his foregroundings were in both episodes *without* his explicit awareness, and to recognize how much *more* powerful and purposeful these foregroundings could be *with* his explicit awareness. Such awareness might also facilitate teachers' shifting among aspects more responsively in the course of authentic disciplinary practice with students, demonstrating a sort of flexibility that might represent another avenue of growth for Mr. S and others.

Implications for Research

In terms of research, it would be beneficial to understand more about the impact different disciplinary foregroundings have on what students come to see as authentic disciplinary activity. For instance, shortly after Mr. S recapped Diane's idea as having to do with speed in the first key drop episode, a visiting member of our research team asked, "Folks, did you hear that reasoning?" A student responded, "Yes, it's based on speed," suggesting that Mr. S's foregrounding of causal factors may have been picked up by students as a sufficient explanation. Exploring potential connections like this between teachers' foregroundings and students' senses of the discipline is a ripe area for future research, with important implications for what students learn through engagement in discussions at the intersection of their ideas and disciplinary practices.

Future research could also target how explicit professional development discussions of various disciplinary aspects impact teachers' classroom practice. Do teachers exhibit enhanced metaawareness about what they are attending and responding to within students' ideas? If so, do they demonstrate more or less flexibility in what they foreground, and to what ends? Such questions could be explored in continuing professional development projects aimed at enhancing teachers' attention and responsiveness to the substance of student thinking.

Endnotes

(1) These statistics come from publicly available 2009-2010 demographic data, not directly cited to protect the anonymity of the school.

(2) That said, there are certainly situations in which foregrounding the identification of relevant causal factors is appropriate, like when engaging in experimental design (e.g., Ford, 2005; Toth, Klahr, & Chen, 2000). The identification of relevant causal factors for a given phenomenon provides useful insights about the phenomenon and predictive power with respect to similar phenomena, and is a publishable finding in various scientific disciplines, such as ecology, epidemiology, etc.

References

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal*, *93*(4), 373–397.

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*, 26–55.

Brodie, K. (2011). Working with learners' mathematical thinking: Towards a language of description for changing pedagogy. *Teaching and Teacher Education*, *27*, 174–186.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*(4), 499–531.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*, 175–218.
Crespo, S. (2000). Seeing more than right and wrong answers: Prospective teachers' interpretations of students' mathematical work. *Journal of Mathematics Teacher Education*, *3*, 155–181.
Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, *4*(1), 37–73.
Empson, S. B., & Jacobs, V. R. (2008). Learning to listen to children's mathematics. In D. Tirosh & T. Wood (Eds.), *Tools and processes in mathematics teacher education* (pp. 257–281). Rotterdam, NL: Sense Publishers.
Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education*, *27*(4), 403–434.
Ford, M. J. (2005). The game, the pieces, and the players: Generative resources from two instructional portrayals of experimentation. *Journal of the Learning Sciences*, *14*(4), 449–487.
Franke, M. L., Carpenter, T., Fennema, E., Ansell, E., & Behrend, J. (1998). Understanding teachers' self-sustaining, generative change in the context of professional development. *Teaching and Teacher Education*, *14*(1), 67–80.
Goldsmith, L. T., & Seago, N. (2011). Using classroom artifacts to focus teachers' noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 169–187). New York, NY: Routledge.
Hammer, D., & van Zee, E. (2006). *Seeing the science in children's thinking: Case studies of student inquiry in physical science*. Portsmouth, NH: Heinemann.
Jacobs, V. R., Lamb, L. L. C., Philipp, R. A., & Schappelle, B. P. (2011). Deciding how to respond on the basis of children's understandings. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 97–116). New York, NY: Routledge.
Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, *7*, 203–235.
Levin, D. M., & Richards, J. (2011). Learning to attend to the substance of students' thinking in science. *Science Educator*, *20*(2), 1-11.
O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology & Education Quarterly*, *24*(4), 318–335.
Pierson, J. L. (2008). *The relationship between patterns of classroom discourse and mathematics learning* (Doctoral dissertation). University of Texas at Austin.
Richards, J. (2013). *Exploring what stabilizes teachers' attention and responsiveness to the substance of students' scientific thinking in the classroom* (Doctoral dissertation). University of Maryland at College Park.
Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3), 499–525.
Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4, Part 1), 696–735.
Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, *12*(1), 5–51.
Schegloff, E. A., & Sacks, H. (1999). Opening up closings. In A. Jaworski & N. Coupland (Eds.), *The discourse reader* (pp. 263–274). London, England: Routledge.
Stivers, T., & Sidnell, J. (2005). Introduction: Multimodal interaction. *Semiotica*, *156*(1/4), 1–20.
Toth, E. E., Klahr, D., & Chen, Z. (2000). Bridging research and practice: A cognitively based classroom intervention for teaching experimentation skills to elementary school children. *Cognition and Instruction*, *18*(4), 423–459.
van Es, E. A. (2011). A framework for learning to notice student thinking. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134–151). New York, NY: Routledge.
Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, *96*, 878–903.

Acknowledgments

This work was supported by funding from NSF DRL-0733613 and NSF EHR/DUE-0831970. All findings, opinions, and recommendations expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation. Special thanks to David Hammer, Paul Hutchison, Eric Kuo, Amy Robertson, and ICLS reviewers for feedback on earlier versions of this work.

Spatial Practices in CSCL Discussions

Benzi Slakmon and Baruch B. Schwarz, The Hebrew University of Jerusalem,
School of Education, Mount Scopus, Jerusalem, 91905, Israel
Email: benzion.slakmon@mail.huji.ac.il, baruch.schwarz@mail.huji.ac.il

Abstract: In CSCL environments, space itself − and not only utterances − are objectified. New questions emerge from the relation between objectified space and student activity: How do students react to the spatial resource opened at their disposal? What are their spatial modes of acting within it? What bearing does the spatial organization have on talk? This paper revitalizes the neglected semiotic connection between language and the city. The literature of the perspective of everyday life and other contemporary spatial literature are reviewed in order to better understand the use of space in CSCL discussions, the ways in which it is socially produced, reclaimed, planned, maintained and used. 37 small group Argunaut discussions were analyzed and examined for their spatial development. Four spatial practices are described. The practice of distribution is unfolded. This study argues that issues of proprietorship play a major role in the way discussions evolve.

Introduction

Temporality, sequentiality, segments, events, turns, episodes: When we think of the conceptual category that holds our thinking about discursive phenomena, we primarily think of *time* (Mehan, 1985). Talk is not usually perceived as a spatially organized phenomenon. As talk is moves into CSCL environments, however, words and utterances are objectified, thus becoming an object for reflection. But it is not only the linguistic plane that is objectified; Space itself-the linguistic platform- is also objectified. As a result, the notion of conversational space gets a new actual meaning. The objectification of the conversational space in CSCL raises new kinds of questions: how do students react to the spatial resource opened at their disposal? What will be their spatial modes of acting within it and what bearing will the spatial organization have on talk itself? By deliberating on the emergence of a spatial order in CSCL discussions, we might be able to better understand the processes involved in the emergence and power of the ideas of the private and the public.

In order to understand the use of space in CSCL discussions − the ways in which it is socially produced, reclaimed, planned, maintained and used− the spatial development of 37 Argunaut discussions were analyzed and examined. These discussions comprise the entire corpus of CSCL work done in an eighth grade classroom during a year-long course in humanities. Because of the nature of the corpus, we were able to trace not only the general scheme of spatial arrangement, but also the exact trajectories of change and transitions in which discursive/spatial practices were introduced and later endorsed. Using conversation analysis alongside the spatial analysis, we could identify the key moments of emergence of new spatial practices, and of norming and changes in the meaning of the practice along the corpus. In the following section we will try to construct the foundations for a discussion on the notion of conversation as seen from the perspective of the spatial, as well as from the perspective of property. We will argue that neglecting these perspectives when analyzing student's CSCL talk conceals certain of its attributes. Moreover, we claim that by enriching the analysis, we are also reconnecting and bringing back the intellectual achievements of the recent literature on space, which was based on the metaphor of language and the fruitful comparison between language and the city.

The Space Density Problem

At times when the discussion is productive and many contributions are being made, it does not take long before the problem of density arises. The contributions rapidly accumulate. The argunaut is programed in a way that gives the students the freedom to enlarge its boundaries. Students can drag their contributions to the edge of the conversational space, width and length, and by doing so to enlarge it. As they do so, they cannot see all of the contributions at the same time the density of the discussion affects the quality of the talk because one cannot distinguish the newer contributions from the older ones, and cannot deal with the content and relational overload. The responsiveness of dialogue is at stake when it is not clear who refers to what, and when. In order to deal with the problem of density, the students developed two spatial practices. The first is *the enlargement of the conversational space*, as can be seen at the discussion showed at the upper right corner of figure 2. The second practice is *'the shout'* (figure 1). Shout takes place when a discussant is placing a contribution randomly on the conversational space, and then *moving it to the center of the space and enlarging it disproportionally*, thus squashing all that is present beneath it. The over-sized comment is held there to be seen for an average of 8-15 seconds and subsequently scaled down and dragged to unoccupied territory. In a dense space governed by no conventions of order, some students find this to be the best way to be heard and recognized, even when it means squashing others. When shouting is practiced in discussions, the students who do not actively make use of it are given less recognition.

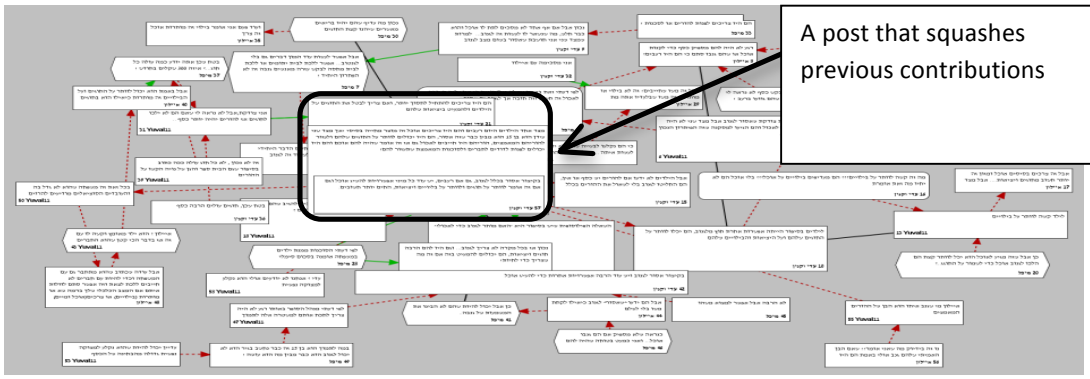


Figure 1. Congestion and squashing in an over-loaded space.

Theoretical Framework

Space's Multiple Meanings and the Learning Sciences

There is a persistent use of the notion of 'space' in the learning science, and particularly in research on CSCL (Erduran et al., 2004; Mercer and Littleton, 2007; Wegerif, 2007; Chin and Osborne, 2010; Schwarz and Asterhan, 2011). Seen from the history of ideas perspective, the term space holds multiple, sometimes contradictory, meanings (Casey, 1997). There are two distinctive ways to use the term: space as entity and space as order. Applied to the textual sphere, the matrix is more complex: using Zoran's (1997) *textual spaciality theory*, the conjunction between these two uses and the two textual levels- the linguistic functions of the *utterance level* and the represented reality of the *world level*- yields a structure with four possibilities for implementing the concept of space in the text: (1) Space as order, applied to the utterance level− that is, on paradigmatic relations and connections between disconnected textual units (2) Space as entity, applied to the utterance level that is the physical, graphic existence of the textual signs (3) Space as order, applied to the world level, which is the entire space of the world: the physical, ideational, psychological (4) Space as entity, applied to the world level. The physical structure of the world as a site of occurrences, plots, and acts. Analyzing the Learning sciences' literature, based on Zoran's Theory it is clear that the term space is used only to describe order in the utterance level. Here we will focus on space as entity.

Talk and the Everyday Life Perspective

The pedagogical challenge of designing talk similarly to that of city planning, is to embrace plurality, in such a way that will lead to generations of forms and practices which will be endorsed and will contribute to the quality and richness of the discursive experience. The idea of the city as a metaphor for language was introduced by Wittgenstein in a famous aphorism (2001, 18): "Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight regular streets and uniform houses". The similarities between speech and urban architecture, and the use of speech as a metaphor for the city, were a generating force in the development of spatial knowledge. Certeau (1984) capitalized on the similarity when he described *walking as a space of saying*. Everyday life critique (Lefebvre, 1991) might be instrumental in the analysis of the ways in which students discursively and spatially act, in classrooms and in conversations, in order to redefine the meanings of spaces through the practices of recurrence, acquiring, reclaiming, distributing, and appropriating it. Everyday life is the intersection of the site in which individual practices, and social constructs interact (Lefebvre, 1991). The other way also applies: the space of interaction between individual practices and social norms is a site of everyday life. According to Gardiner (2000), it is in these sites that we develop our individual and societal capacities, and only in such spaces, do we become full members of the community (Fenster, 2012).

The idea of appropriating space through recurring acts is closely related to Aristotle's definition of 'having'. Aristotle perceived possession as a mediated act (praxis) or movement between the proprietor and the thing (Metaphysics, book V). The Greek word describing possession is hexis. The word can be also described as a qualification that originates in practice and habituation. The physical dimension of the process of hexis is captured in Latin translation as habitus. The term habitus refers not only to the process of acquiring through recurrence, or to the habitual aspect of being accustomed and fluent in the action, but also to the growing capability, readiness and ease to perform the act in the same, but also changed, environments. The experience of ownership- of the conversational space- is the accumulated outcome of recurring bodily practices of spatial doings. The knowledge produced in the process of 'having' is therefore, localized and privatized. It should be highlighted that hexis, due to its habitual and bodily nature, is the cornerstone of Aristotle's ethical virtue, a virtue that neither can be appropriated through direct teaching nor developed naturally in man.

Setting and Context

The data presented is taken from a design-based study, conducted as a year-long humanities course in an eighth grade classroom of 13 girls and 15 boys. The course met six hours per week, and was taught by the first author of the article. The goal of the humanities course was to develop humanistic curricula based on an elaborated form of a *community of philosophical inquiry* (Lipman et al. 1980), in order to engage students in learning to dialogize. The course was divided into three learning cycles, and dealt with ethical and epistemological issues. Neither argumentation nor rigorous school of philosophical thought was addressed directly, as the course stemmed to some extent from the questions of the students and from their interests (Slakmon & Schwarz, in press). Since objectifying talk was one of the major goals for students to become more reflective about both talk and thinking, throughout the third learning cycle, we added a weekly group discussion in the Argunaut system.

The Argunaut system is a CSCL tool designed for promoting synchronous argumentative discussions. Based on the Digalo system (Schwarz and de Groot, 2007), Argunaut appends a monitoring and intervention unit to Digalo. Argunaut has emerged from the need to moderate multiple synchronous discussions in classroom. It provides the moderator with public and private communication channels with the discussants and 'awareness tools' for monitoring ongoing discussions (Schwarz and Asterhan, 2011). The moderator builds the conversational space in advance, adding to it the question/issue to be discussed, and the argumentative/communicative ontology he/she wishes to introduce. The moderator builds numerous conversational spaces in parallel and assigns each predetermined group to its designated conversational space. By doing so, the group finds the question/issue and the ontology in the tool bar. The ontology becomes a speaking/thinking tool since discussants need to select an argumentative category (represented as a distinct geometrical form). Furthermore, students are expected to link their posts to other posts by using one of the communicative connections ('arrows') present in the ontology (green arrow for agreeing, red for criticizing, black for neutral reference). The paper deals with the question of how they act in the conversational space, not from the point of view of what they post, but from the point of view of managing the conversational space. When students engage in an Argunaut conversation, they need to address the question of spatial organization. Like chats, the utterances students produce in the Argunaut system are fixed and do not disappear from conversational space over time. But unlike chat, there is no restraint channeling discussion into threads, so whenever a student posts a contribution on the conversational space, several decisions must be made to differentiate the Argunaut way of communicating from more ordinary forms. A student must decide where to place a contribution and also which of the previous comments to connected with arrows to the utterance. As shown in figure one, as the conversation develops, it becomes overloaded with utterances, and as a result, students find it difficult to find their way in it. Nothing distinguishes earlier entries from later entries so it becomes increasingly difficult to distinguish more important from less important utterances. The CSILE research team faced similar problems and came up with the dialectic function of "rise above" (Scardamalia, 2004). Argunaut does not include any tool for envisioning convergence, leaving the issue for the students to deal with. It became increasingly difficult to distinguish the important from the less important, thus hurting the potential of referentiality, meaning the ability to base an utterance on previous, but not immediately preceding utterances. Since cultural and discursive achievements are based on the accumulation and preservation of traditions, this state of affairs prevents rich and cultured construction.

Eight Argunaut session days were conducted in groups consisting of 3-6 students. Weekly discussions were held with the exception of the first two sessions that took place during the first learning cycle and were held chronologically four months earlier to the third discussion. A total of 37 discussions were produced. All of them were spatially analyzed. Argunaut sessions were held at the school's computer lab and in each session, up to three groups worked simultaneously with the Argunaut, while the other groups were assigned to different tasks, outside of the computer laboratory. Each group formed its own original discussion map. The groups were formed by the teacher, whose attempts to maintain fixed groups each week, were only partially successful, due to student absenteeism. Students were familiar with the computer lab since they had used it all year. The following analysis is based on analyzing the groups' trajectories of participation, and more precisely, collective spatial usage trajectories of the entire 37 discussions.

Order of Occurrence I: First Glimmer of Spatial Order

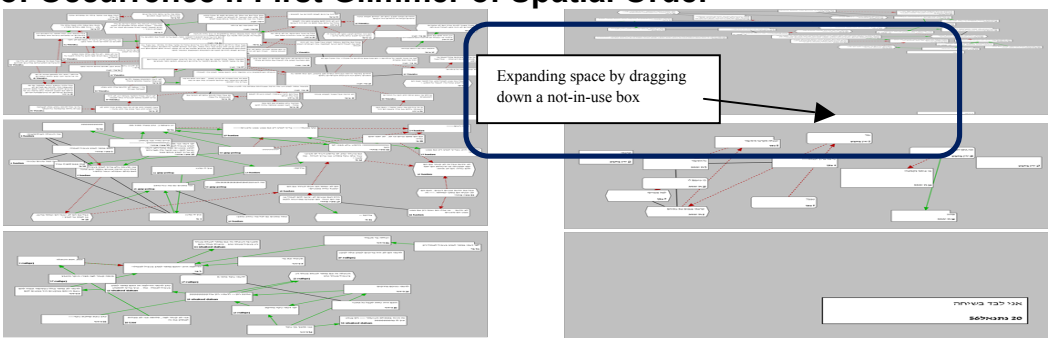


Figure 2. Maps of the six Argunaut discussions held in the fourth discussion day (discussions no.12-17). On the upper right: the practice of enlargement

In all of the 19 first discussions, students were posting their contributions in the conversational space wherever they wanted, and without either consideration for issues of private/public spaces, or any indication of a time-based organization. Figure 2 show how the six discussion maps of the groups looked at the end of discussions no. 12-18. All previous discussion looked the same with regard to the scattering of the contributions. But at discussion day 5, at the following week, only three discussions were held (18-20, figure 3) because some students were absent. As seen, the discussion maps spatial order was as accidental as it was before, except in discussion 20, held by Shira, Dor and Sharon (the upper map in figure 3). Thoughtful spatial organization appears in discussion 20 for the first time. In it, the right column consists of all of Dor's contributions. The cluster located at the bottom of the center of the map consists of Sharon's contributions, and the left column consists of Shira's. A blurred spatial order emerges in this discussion, but some questions need to be asked: How did the order emerge in the first place? What is its function? How do the students comprehend and interpret it? And finally, what will become of it in the following lessons? All questions are temporal by nature, therefore appropriately addressed by trajectory analysis.

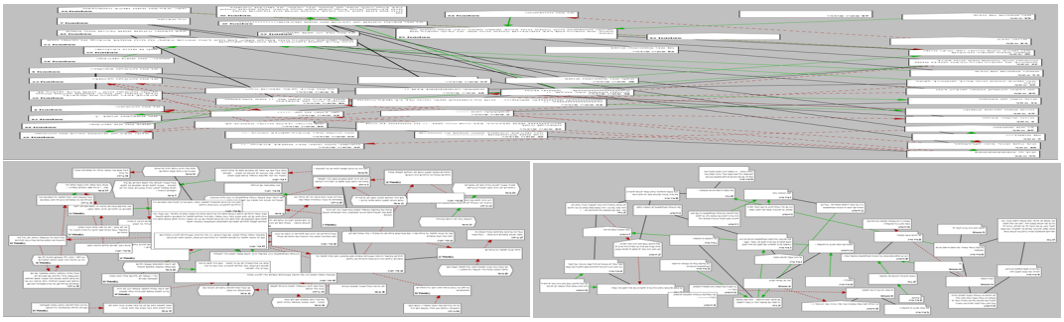


Figure 3. Maps of the three discussions held in the fifth discussion day (no.18-20). Discussion no. 20 highlighted

The students were dealing with a moral dilemma posed by a classmate in a story he composed. After reading the story, the participants were assigned to the Argunaut to discuss the problem. Sharon, Dor, and Shira, the three students involved in the discussion (20), placed their contributions in the center of the space, in no observable order. Eight and a half minutes after they started the conversation, Sharon used *shouting* to post the following: "pputttttttt a little bittttt of orderrrrrrr in the dialogue we're havingggg…!! This is theeeeee perfect dialogue of ours". But her contribution did not receive any obvious response. The discussion continued. Shira formulated a well-informed argument and received Dor's endorsement. They each placed their contribution wherever they found space. Three minutes after Sharon's call, the flow stopped. At this point, Shira's contributions were scattered all over the space. Among them were comments 5 and 10, positioned on the left side of the space, one above the other. Shira created a gap between them and moved one of her previous contributions between them. The next thing she did was to align the newly created three box column by moving the upper box. The entire move took six seconds. It took six seconds for Dor to detect the move. He dragged one of his contributions to the extreme right, opposite of Shira's column. There Dor used contribution as the cornerstone of his column. He continued to rapidly build his column from his previous comments. Sharon started organizing her eight contributions just as Dor and Shira finished organizing theirs. The process of organization took place in a flicker of less than two minutes. In it, no new utterances were produced; each column was made of contributions authored by a single participant. The placement of the contributions inside the columns was driven by a spatial-organizational logic: *The columns have no chronological consistency*; we

can see that with regard to all three discussants, there is no connection between the production time of the contributions and their place in the column. It is important to note that this time the organization of the space emerged only *after* the discussion was almost entirely completed. It did not emerge as an integral component of the development of content and ideas. The session lasted for another 12.5 minutes. As seen in figure 4, an open, un-utilized, space was created between Dor's and Sharon's columns, and another free space was found above Dor's column. In the remaining minutes of the discussion both areas functioned as *public spheres* which became the place for introducing new comments and exchanging ideas. New contributions were posted and placed first in it, sometimes accompanied with a *shout*, and then dragged to the *private sphere* of the contributor, as if it were a conversational capital belonging to the producer.

Order of Occurrence II: Recognizing the Practice of Distribution

Five discussions were held the following week (21-25), at day 6. Both Shira and Dor regrouped to different groups with other classmates. Sharon did not attend that day. As seen from figure 5, Dor (discussion no. 22) duplicated the practice of distribution, whereas Shira's discussion (no. 21, figure 4) was generated without any other spatial organization than shouting. Although Shira was the first to introduce distribution, this time she posted and placed her contributions randomly, like everybody else. Consequently, the practice of distribution was neither introduced to the other participants nor reinvented. In Dor's (22), we see the *migration of the idea of spatial organization through the practice of distribution* for the first time in the entire corpus. Dor's session ended with spatial organization very similar to the one introduced for the first time a week earlier in his discussion (20). Not only did Dor import the practice of distribution, but he also performed it fervently keeping his own 'property' on the right-hand side of the space, just as in the original discussion. As in the previous week, distribution and privatization only occurred after sixteen minutes, and yet again, Dor did this while ignoring the chronological aspect of placing his contributions. He acted swiftly, and all other participants followed his example, and made their own columns. Although Shira does not practice distribution, one practice Shira did reuse was shouting. Reviewing all of previous and parallel discussion in this corpus, we traced only two shouts prior to discussion 20; that is, the practice had no contagious effect. In fact, it was not recognized as a practice: in both two first shouting cases, the gesture did not repeat itself by the shouters, nor was it reenacted by their peers. But in Shira's discussion group (22), not only did it occur, but it was also repeated over and over again, by all three participants. The practice was recognized and endorsed immediately after Shira introduced it for the first time and became a *common and shared* practice, a real hit: twenty three shouts appeared in the discussion. In the social sense then, shouting was generated in this session.



Figure 4. Shira's group sixth day discussion (no.21, left) and Dor's group sixth day discussion (no. 22, right)

Order of Occurrence III: Endorsement

Figure 5 shows the ways discussions ended a week later. Between the sessions the class conducted a teacher-led reflective lesson, in which Dor's way of spatial organization was introduced by the teacher in the presence of his classmates (the practice of shouting was not discussed by the teacher).However, the teacher *incorrectly* referred to it *not only as spatial but also as chronological* (as if the upper boxes are the first to be created and so on). And he praised the practice suggesting that this might contribute to student's ability to reconnect ongoing discussion with adjacent contributions, and to enrich the conversation and make it more reflexive and cohesive. The following week, the teacher's incorrect assumptions turned out to be productive: four out of the five discussion groups practiced *spatial and chronological distribution*. For the first time in the corpus, and simultaneously in all four groups practicing distribution, distribution was the organizing logic of space and time right from the beginning. The contributions were posted in accordance with it. Surprisingly, the only group to act differently was Shira's, the unaware inventor of the norm (at the upper left part of figure 5). Spatially, her discussion group started with random order, reorganized into columns and scattered again before finally becoming organized. The loosely maintained spatial organization in this group was not accompanied by a chronological one.

Small-group modification of a social norm might be seen as a different, and maybe higher, level of endorsement and mastery. This is the case presented in the discussion held by Nadav, Amos, and Adam (at the lower left part of figure 5). As the session ended, we find the discussion map to be organized with two columns at the poles of the space and a boxes-made robot in the center. They started the conversation adhering to the emerging norms, and at one point near the end of the session they transcended them through play, as they started to refer to the boxes as imaginative building blocks.

Order of Occurrence IV: Consolidating the Practice of Distribution

Once repetition of social action is played out by different agents and identified as such, one might presume mastery in performing the practice (Sfard, 2008). This is the case shown here with regard to the spatial and chronological organization and with regard to the act of shouting. The final Argunaut session held two weeks afterwards. In between, the teacher did not refer to the topic and no further Argunaut session was held. In the final day, all four discussion groups acted in accordance with the norms of spatial and chronological organization. More interesting is the fact that only one group actually started with the order right from the beginning. In all three other cases, sessions began with random order up until a point where one participant called for order. *Without any further talk regarding the 'what and how'*, all discussions became involved in a short period of massive reorganization of space, and all three discussions reorganized in accordance with the spatial norm endorsed. From this point on, all of the groups started posting in chronological order as well. The endorsement in these three discussions is a *social-action, based upon individual's reference to the same activity-substructure*.

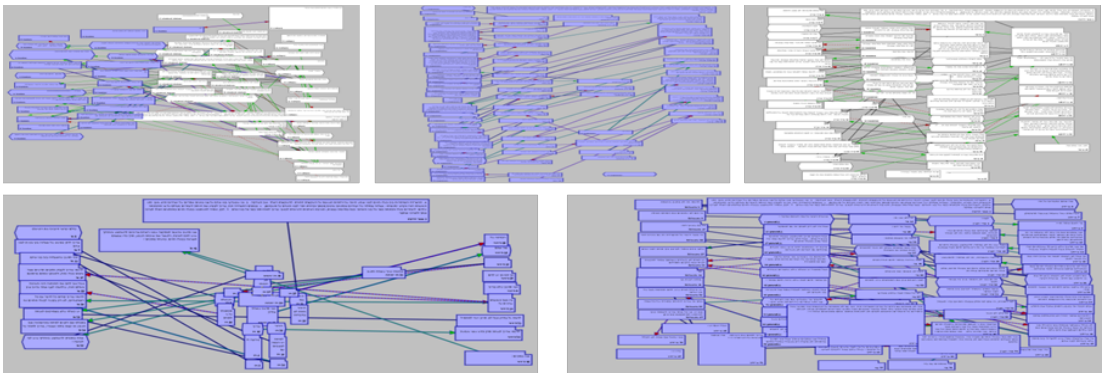


Figure 5. Group's discussion maps at the end of the seventh day (discussions no. 26-30)

Discussion

The question of ownership in conversational spaces (Slakmon and Schwarz, in press), which has been neglected from the analysis of CSCL and peer discussion, turned out to be a major factor in students' discursive behavior. Before the endorsement of the norm of distribution, *all* students treated the conversational space as public space. The question of possession did not arise at all. Everyone could have posted wherever suitable, even with squashing the other's contributions. As a result, the students' voices were not heard, and they were subject to constant squashing. There were no complaints about the squashing, but the discussions evolved with no reference to the past, that is, its latter parts. Therefore, it lost the potential it had to establish traditions, in the sense of building on each other's ideas, accumulating socially acknowledged insights and intellectual achievements.

Both squashing and shouting are egoistic solutions, and only contributed to worsening the problem, since they delete past contributions in order to be heard. Both solutions were destined to fail: the previous parts of the conversation that function as a background and as a context are preconditions that must exist if one's voice is to be heard. In this sense, all students treated the conversational space as owned by none of them, and it suffered from the classic *tragedy of the commons* (Hardin, 1968). Only after the emergence of distribution, and the introduction of the idea of ownership, did the space turned out to be part of the public sphere. Space was socially produced in the collaborative act of planning. The significant collaborative achievement of students here is realized when we see planning as an act of arranging space in accordance with the goals and principles of the power holders (Fenster, 2012). In this sense, when students are involved in the social production of space, it can be inferred that they act within it as power holders. If not, their entire discursive strategies would have been different: opposing, accommodating or subversive. Moreover, the practice of dividing space into private and public led to the decline of the earlier strategies aimed at solving the problem of the density of the space*: squashing and shouting only existed in the discussions where the students did not divide the conversational space.*

The sense of ownership has a significant effect on the way students hold discussions in the conversational space, In the public sphere in which the ritual of posting a comment awaits others to read it and

then, moving it to one's own space. This only happened when students had private space to return to. They guarded their private space and the public *at the same time* and the private and the public spaces nourished and strengthened each other. The space was divided into private and public spheres on the basis of a previous agreement on perceiving it as a collective source. Through the joint production and maintenance the private and the public, discussion thrives.

Interesting questions regarding spatial practices are answered in this paper but need further analysis, among them, the analysis of the in-groups politics of endorsement, and the question of the relations between spatial practices and discussion quality. With regard to the later, initial analyses suggest that discussions spatially organized to private and public spaces were more productive in terms of number of contributions and of contributions that functioned as knots of convergence, or of joint reference by the entire group.

Students gave up the privileged freedom of posting anywhere, and opted instead for the more restricted arrangement, regulated by the social norm of distribution. The ability to co-exist became a procedural norm and no longer was it left to the jurisdiction of each participant. This freed the students from the need to decide according to which criteria, if any, they should squash other contributions. The mere act of being in the conversational space had been ensured, and was no longer a matter of treaties, friendship, time or content. A tolerant way of being together at the conversational space gained prominence, and in so doing the students developed an enhanced way of dealing with heterogeneity without muting each other's voices.

References

Aristotle. (1971). *Metaphysics (books gamma, Delta and Epsilon*). Oxford : Clarendon Press.

Certeau, M. D. (1984). *The Practice of Everyday Life.* Berkeley, CA : University of California Press.

Chin, C., & Osborne, J. (2010). Supporting Argumentation Through Students' Questions: Case Studies in Science Classrooms. *Journal of the Learning Sciences*, 19(2), 230 - 284.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6), 915-933.

Fenster, T. (2012). *Of Who Is This City? Tel-aviv: Hakibutz Hameuchad (Hebrew).*

Gardiner, M. E. (2000). *Critiques of Everyday Life*. New York and London: Routledge.

Hardin, G. (1968). The Tragedy of the Commons, *Science*, 13, Vol. 162, 1243-1248.

Lefebvre, H. (1991/1967). *The production of space. Oxford,UK : Blackwell.*

Lipman, M., Sharp, A. M., & Oscanyan, F. (1980). *Philosophy in the classroom.* Philadelphia : Temple University Press, (2nd ed).

Mehan, H. (1985). The structure of classroom discourse. In T. A. van Dijk (ed.), *Handbook of Discourse Analysis (Vol.3)* (pp.115-131). London: Academic Press.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: a sociocultural approach.* London, UK: Routledge.

Scardamalia, M. (2004). CSILE/Knowledge Forum®. In Education and technology: An encyclopedia(pp. 183-192). Santa Barbara: ABC-CLIO.Schwarz, B. B. & De Groot, R. (2007). Argumentation in a changing world. *The International Journal of Computer-Supported Collaborative Learning, 2*(2-3), 297-313.

Schwarz, B. B., & Asterhan, C. S. C. (2011). E-moderation of synchronous discussions in educational settings: A nascent practice. *Journal of the Learning Sciences*, 20(3), 395-442.

Sfard, A. (2008). *Thinking as communicating.* Cambridge : Cambridge University Press.

Slakmon, B. & Schwarz B. B. (in press). Disengaged students and dialogic learning: the role of CSCL affordances. To appear in: *International Journal of Computer- Supported Collaborative Learning*

Wegerif, R. (2007). *Dialogic, Educational and Technology: Expanding the Space of Learning*. New York: Springer-Verlag.

Wittgenstein, L. (2001/1953). *Philosophical investigations*. Oxford, UK : Blackwell, (3rd ed).

Zoran, G. (1997). *Text, world, Space.* Tel-Aviv: Hakibutz Hameuchad (Hebrew).

Collective Immersive Simulations: A New Approach to Learning and Instruction of Complex Biology Topics

Michelle Lui and James D. Slotta,
University of Toronto, 252 Bloor Street West Toronto, Canada, ON,
Email: michelle.lui@utoronto.ca, jslotta@oise.utoronto.ca

Abstract: This paper presents the design of an immersive simulation and collective inquiry activity for exploring evolutionary concepts in a Grade 11 Biology course. Researchers and a high school science teacher co-designed a curriculum around a room-sized simulation of a rainforest. Using several large displays stitched together on each wall of the room, we created an immersive rainforest environment in which students worked collaboratively as "field researchers" to observe changes in life forms over two hundred million years and gathered evidence of evolution. The complex sequence of student interactions within the EvoRoom environment, as well as all materials, including large immersive displays, aggregated visualizations, and tablet applications were carefully designed as short inquiry activities that complemented the broader curriculum. This paper presents our designs over two iterations in terms of several key features that enhanced students' collective immersive experience and learning of evolutionary biology.

Introduction

Educators and researchers have long struggled to help students achieve deep understanding of complex science concepts, and to help students refine reasoning and communication skills, such as critical thinking and collaboration (NSF Taskforce for Cyberlearning, 2008). In high school biology, concepts of evolution and biodiversity are notoriously challenging, due in part to their complex systemic nature (Slotta & Chi, 2006), their multidisciplinary nature (e.g., genetics, biogeography, paleontology), as well as students' incoming ideas, which are often inconsistent with the scientific theory (Demastes, Good, & Peebles, 1995; Mayr, 2001).

Inquiry-based learning has been advanced as an instructional approach where students are encouraged to develop deep understandings and scientific reasoning, by emphasizing the posing of questions, collection and analysis of data, and construction of evidence-based arguments. Inquiry-based learning has shown promise for teaching evolution, as exemplified by projects like the Biology Guided Inquiry Learning Environment (BGuILE; Reiser et al., 2001) and GenScope™ (Horwitz et al., 1998).

In recent years, researchers have begun to reconsider the role of the physical learning environment and to experiment with augmenting learning activities in digitally augmented physical spaces (i.e., mixed-reality environments). These spaces offer new ways of engaging groups of co-located students with abstract science concepts that have traditionally been taught or addressed through more didactic forms of instruction (Price & Rogers, 2004). Moreover, such physical learning spaces, when used in combination with inquiry-based learning activities, have shown positive outcomes in facilitating creativity and reflection (e.g., Facer et al., 2004; Rogers & Muller, 2006) – offering a more hands-on approach compared to traditional inquiry-based learning where students typically work autonomously as individuals or pairs on single machines (Slotta & Linn, 2009).

Collective Inquiry and Smart Classrooms

The present study seeks to leverage technology-enhanced learning environments in support of more complex and participatory forms of scientific inquiry that engages co-located students in an investigation about evolution. Our goal is to create a comprehensive curriculum that embodies a pedagogical perspective known as Knowledge Community and Inquiry (KCI), where students are supported to work as a collective scientific body, creating a knowledge base and using it as a resource for subsequent inquiry (Slotta & Najafi, 2012).

Transforming classrooms into "knowledge communities" can engage students in more authentic scientific inquiry (Brown & Campione, 1990), for example with small groups of students working together like research teams within a broader scholarly community to jointly negotiate issues of a shared problem. By generating and building upon each other's ideas, students take greater responsibility for ultimately fostering their own understanding. In previous KCI studies, wikis were used to support knowledge communities, with findings showing positive correlation between students' contributions to collaborative inquiry and their achievements in the curriculum (Peters & Slotta, 2010). However, much like traditional inquiry-based learning, communications tended to be asynchronous and distributed, with students mostly working on their personal computers. To this end, we seek to engage students in collective inquiry as a knowledge community about evolutionary biology within a "smart classroom" environment, where the physical environment is intertwined with a set of digital tools and materials to scaffold seamless and dynamic collaboration and real-time face-to-face interactions while capturing the collective wisdom of the entire class (Slotta, 2010).

 © ISLS

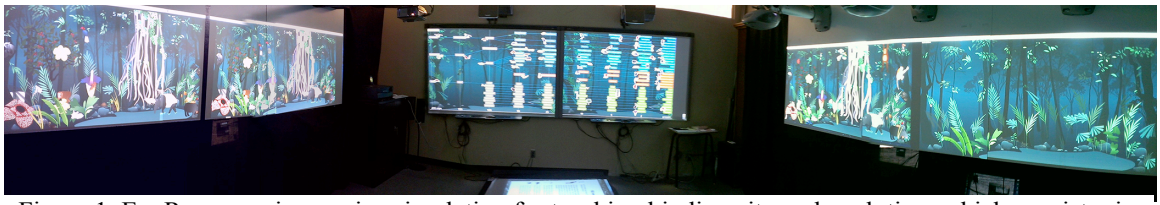


Figure 1. EvoRoom, an immersive simulation for teaching biodiversity and evolution, which consists six projected displays (three on each side) and two interactive whiteboards (middle).

Immersive Simulations

Inspired by the research tradition in immersive virtual worlds, such as River City (Dede, 2009) and Second Life, we are investigating a possible new educative role for immersive simulations, where the smart classroom is converted into a rich simulation, and conceptual content is embedded in ubiquitous technology to support co-located students in learning as a community (Figure 1). The immersive, room-sized environment is responsive to student observations (recorded via tablet devices), with real-time emergent visualizations that serve to capture and aggregate student observations for purposes of knowledge building and discourse. Our research is concerned with designing inquiry activities that complement and help to define such immersive environments, where students are engaged as a whole class, jointly negotiating problems and working towards a common goal. In this novel form of inquiry-based learning, called collective inquiry, students are encouraged to think deeply about materials and develop their own understandings, but with an emphasis on collective knowledge or progress over individual understandings (Peters & Slotta, 2010; Slotta, 2010).

Background & Related Work

Participatory Learning and Physical Digital Spaces

Wilensky and Resnick (1999) pioneered the use of embedded, ubiquitous computational media to support science learning, including the use of role-playing activities and non-desktop technologies. This is illustrated by the concept of participatory simulations, in which students themselves serve as the elements of a simulation (Colella, 2000). For example, Colella (2000) transformed students into potential virus carriers through wearable computers, with the mission of greeting as many peers as they could without becoming "sick". Inspired by participatory simulations, another approach called Embedded Phenomena (EP; Moher, 2006) features a persistent scientific simulation that is "embedded" into the walls or floor. Students are tasked with discovering, monitoring and manipulating the state of the simulation and gathering evidence in the course of their inquiry about the phenomenon. Other examples of digitally augmented physical learning spaces include SMALLab, a room with digitally enhanced walls, floors and interactive technologies that supports new forms of student inquiry. For example, high school students studied geologic evolution by collaboratively constructing and monitoring the Earth's crust, identifying uplift and erosion over time (Birchfield & Megowan-Romanowicz, 2009). Using various input devices (e.g., glowballs, Wii remotes, wireless game pads) and a projected interface on the ground, groups of students were responsible for building, maintaining or evaluating a cycle of the geologic clock. The intervention resulted in significant achievement gains, demonstrating the promise for further research regarding face-to-face interactions in a computationally augmented physical space, and distributed roles through a generative process that unfolds over time.

Our research is motivated by these projects, extending the role of immersive participatory simulations into a more coherent pedagogical framework, where students are engaged in scientific inquiry as a knowledge community (i.e., collective inquiry), and their experience within the simulation is carefully scripted within a broader curricular design. We designed an immersive simulation for teaching biodiversity and evolution topics in high-school biology courses to understand the following research questions:

- RQ1: How can an immersive simulation be designed to support students in understanding evolutionary biology?
  - 1a: How should an immersive simulation be designed within a broader set of curriculum activities?
  - 1b: What forms of activities and materials support student engagement with immersive media and drive reflection about evolutionary biology?
- RQ2: How can a collective experience within an immersive environment serve to advance a collective epistemology where students come to see their learning as a community effort?
  - 2a: How do we encourage students to respond meaningfully to ideas of their peers, in real-time?
  - 2b: How do we represent community progress and structure its advancement, such that students must engage productively with the aggregated products of their peers' inquiry?

Method

Following a design-based methodology (Brown, 1992; Design-Based Research Collective, 2003), the immersive simulation was designed and evaluated over two iterations as part of a Grade 11 Biology course. Prior to the first classroom enactment, our team of researchers, technology developers, and a high school teacher met regularly for approximately one year to co-design the curriculum activities and the immersive simulation itself. A pilot study prior to the first enactment was conducted to evaluate the immersive environment, including the accessibility of our materials and the relevance of activities to the topic of evolution (Lui & Slotta, 2013).

Participants included students from class sections of Grade 11 Biology taught by our co-design teacher. The first iteration included two sections totaling 45 students aged 14-16. The second trial occurred in the following academic year and included two sections totaling 54 students (aged 14-16). In both trials, a pre- and post-test was administered, and all student observations and notes were collected for analysis. As well, video and audio recordings served to capture patterns of interaction within the EvoRoom environment.

EvoRoom

EvoRoom is an immersive simulation of the rainforest ecosystem of Borneo and Sumatra. Implemented within a "smart classroom" research environment, the room is equipped with computers, servers, projection displays, and customized software to coordinate the flow of participants and content materials, as well as to collect data during the activity. During the collective inquiry activities, students take on the role of "field researchers," working in various group configurations to complete tasks delivered to them on their personal tablet computers. The tablets help to place students in small groups, scaffold their activities, collect observations, and give real-time updates and resources. Student observations and reflections are aggregated and displayed on the interactive whiteboards in real-time.

The collective inquiry activities within the smart classroom were co-designed with the teacher to fit seamlessly within a broader high school biology curriculum, in topics of evolution and biodiversity. Running for approximately 10 to 12-weeks, the integrated curriculum includes in-class activities, homework, a field trip to the zoo, as well as two collective inquiry activities with immersive simulations. One of the collective inquiry activities focuses on the topic of evolution. Students work individually, in small groups, and as a whole-class to gather evidence of evolution by observing changes in life forms within the simulation as it is advanced (by the teacher) across two hundred million years. The second collective inquiry activity focuses on the topic of biodiversity. Prior to the activity, students are to make predictions about how certain environmental factors or changes (e.g., tsunami, earthquake, low rainfall) that occurred within a single season could change the biodiversity over a five-year time span. In the immersive environment, students are presented with four different versions ("scenarios") of the rainforest ecosystem, challenging them to explore the differences between these four rainforests and to locate the scenario that resulted from the variable or factor they explored in their earlier predictions. The present paper focuses on the first EvoRoom activity, on topics of evolution, including the relevant in-class and homework assignments associated with the immersive experience.

Iteration 1: Design and Enactment

At the beginning of the unit, students were assigned to one of four specialist categories (plants & insects, birds, primates, and other mammals), which they held for the duration of the full curriculum. Students were provided with a field guide of a set of species that would appear in the rainforest ecosystem of our immersive simulation. As their pre-activity homework, students wrote a blog post about how their species were related to one another.

Students visited the EvoRoom in groups of between 10-12 students at a time. Within the EvoRoom, these students were split into four groups. Each student was provided a tablet computer for the duration of the activity with a custom designed application that navigated the students through the activity as well as scaffolded students to work together and in collecting data throughout the session. Students visited the smart classroom over two days (for approximately 45 minutes each time). For the first session, students examined the Borneo rainforest as it may have appeared at nine different time periods (i.e., 200, 150, 100, 50, 25, 10, 5, 2 million years ago, and present day). Students were asked to go to each station (from 200 to 2 million years ago) and look for their assigned specialty species as part of a larger team consisting of different specialists. If the species were not present, they were asked to identify their predecessors from a short list that popped up. Their answers were recorded, resulting in the emergence of an aggregated, interactive cladogram (a diagram showing relatedness among species) on the interactive white boards at the front of the room.

In the second session, students entered the rainforest with its "state" set to two million years ago, which approximated Sundaland, a region in Southeast Asia predating Borneo and Sumatra. At this point the teacher used a teacher control tablet to "accelerate time" and showed the resulting geologic events in the Sundaland landscape. Over the span of two million years, sea level changes broke Sundaland's central landmass into a peninsula and several islands, which included Borneo and Sumatra. Setting the room's timeline to present day, one side of the room now showed Borneo's ecosystem, while the other side showed Sumatra's flora and fauna.

Students noted the presence of their assigned species in this new context and in the final step, the students came together in their teams to collectively document evidence of evolutionary differences they "observed" between Borneo and Sumatra (i.e., resulting from their separated state). Students were encouraged to discuss their ideas with others and to post ideas about evolution processes that might have occurred. Their notes were aggregated to the interactive whiteboard, which visibly represented the collective knowledge base of the students at the end of the activity. The teacher was able to use the content of this display, which allowed interactive filtering of the notes by evolutionary concepts and species, to lead a synthesizing discussion.

Iteration 1: Outcomes

Pre/Post-test

The Concept Inventory of Natural Selection (CINS; Anderson, Fisher, & Norman, 2002) was used as a source of the conceptual elements on the pre-/post-assessments. A paired-samples t-test was conducted on the pre- and post-CINS questions to evaluate whether the curriculum supported students in understanding evolution concepts. 33 of 45 students completed both tests. The mean post-test score (M=78.94, SD=15.95) was significantly greater than the mean pre-test score, for CINS items (M=56.34, SD=17.16), $t(32) = 7.14$, $p < 0.001$. Because these items have been developed and validated by assessment researchers as a measure of the evolutionary processes concerned with natural selection, we are satisfied that this overall curriculum engaged students and helped them to learn within this notoriously challenging domain.

Curriculum and Activity Artifacts

To understand how students engaged with various components of EvoRoom we tracked their participation and completion throughout the curriculum. 60% of the students wrote a blog post about how their assigned specialist species were related to one another as their pre-activity homework, while 31% completed their post-activity homework. As a benchmark comparison, 69% of students completed a similar assignment as part of the evolution unit (but not related to EvoRoom) that was graded by the teacher for completion. As part of our analysis of student performance during the collective inquiry activity, we examined how well students did when asked: *Which of the following is most likely [their assigned] organism's ancestor?* Of 1112 answers collected by the student tablets, 78% were correct, with an upward trend of accuracy compared to time period (Figure 2). Since the evolutionary lineages of most organisms present become less ambiguous as the time periods reach closer to present-day, we feel that the students' observation accuracy indicates that students were indeed paying attention to the task at hand and engaging with the media appropriately.

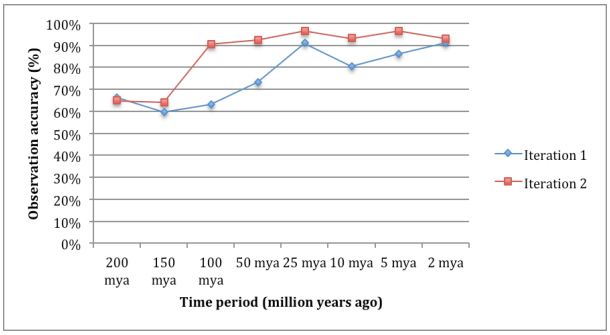


Figure 2. Observation accuracy by time period.

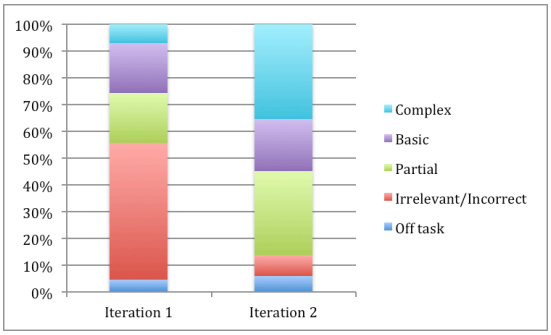


Figure 3. Distribution of evolution explanations' KI scores.

At the end of the activity, students were asked to contribute to the following question: *What evolutionary forces do you think were at play (in this environment)?* Students chose concepts from a predefined list and explained their answers with evidence of their thinking. 43 explanations were collected from the first study which were scored using a 0-5 Knowledge Integration (KI) scale that rewards valid scientific connections between concepts (Linn & Eylon, 2011). The average KI score was 2.7 (SD=1.05) with an average of 33.3 words per explanation (SD=33.28). The distribution of the scores indicates that a large proportion of the notes were scored as Irrelevant/Incorrect (Figure 3).

Design Iteration

Outcomes from our first iteration of the EvoRoom design showed that students significantly improved their understanding about evolution (based on their pre-/post-test results). While the effects cannot be attributed to our EvoRoom alone, these results demonstrated a positive conceptual change when students learned about evolution in a curriculum that included EvoRoom. To further understand how EvoRoom itself impacted

students' ideas about evolution, we needed to revise the pre- and post-test assessment to include deeper conceptual questions targeted to the pedagogical goals of the activities.

In terms of the curriculum activities that preceded students' experience in the immersive environment, we wanted to improve completion rates but did not want to include a competitive mentality such as that evoked by teacher-graded assessments. We sought to encourage a sense of student ownership in the project. Instead of asking students to *review* a field guide, we asked students to collectively *create* the field guide they would be using in subsequent activities. In addition, we wanted students to feel that they had actually had a hand in the design of the immersive environment. Therefore, we revised the curriculum activities that included students researching the Borneo rainforest environment through two hundred million years.

During their collective inquiry activities, students showed considerable focus in their interactions with the materials, but we realized that students had spent more time as individuals in the immersive environment (working towards a collective goal) rather than engaging in the collaboration and deep reflective practices that we had envisioned for our co-located immersive inquiry environment. For example, in the portion of the activity that asked students to locate their assigned organism's evolutionary predecessors, their efforts were displayed in a cladogram that all students in the room effectively created together, but the designed interactions tended to be focused on individual students observing the rainforest "walls" and reviewing materials on their tablet, and then recording their answers. For the second iteration, we sought to improve student-to-student interactions within the immersive environment and while encouraging them to collaborate on more reflective activities.

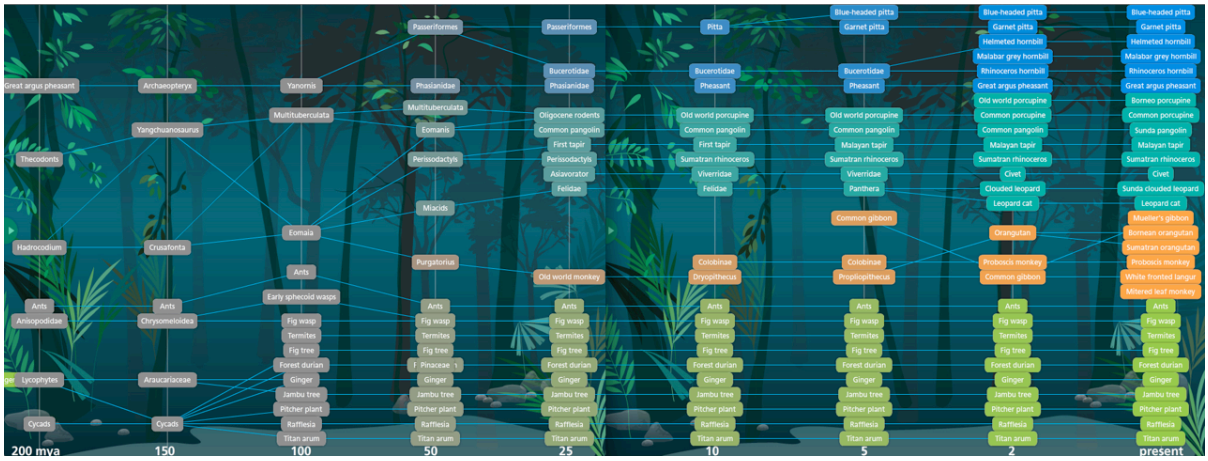


Figure 4. Cladogram created from observations in iteration 2.

Iteration 2: Design and Enactment

The second iteration of the EvoRoom curriculum began with a short lecture delivered by the researchers about the changing nature of science discussing the merits of large scale collaborations (using the Human Genome Project as an example) and were introduced to ideas about themselves working in a knowledge community. Student were again assigned to one of four specialist categories (plants & insects, birds, primates, and other mammals), but was asked to work with other students (across two class sections of Grade 11 Biology) to complete a field guide on a wiki page on their class website. Each student was assigned specific species and categories (e.g., habitat, life cycle, physiology) that they would research. Students were also told that they would collectively create an immersive environment that they would experience together by researching their assigned species' role in the Borneo rainforest over two hundred million years. The premise of the collective inquiry activity, and its organization were similar to that of the first iteration, with a few important changes:

- Two 45-minute sessions were consolidated into one 75-minute session
- When students were asked to identify their assigned species' evolutionary predecessor at stations that displayed the rainforest at different time periods, a student guide (acting like a docent at a museum) was present to help with the task.
- Teams of different species specialists actively compared adjacent time periods (e.g., 200 vs. 150 million years ago) with reflective question prompts. Additional information about the time periods augmenting the rainforest "walls" were provided to students.
- The final step where students discussed and posted ideas about evolutionary processes was further structured to ask students to think about specific species as well as to include artifacts as their evidence. There was also explicit instruction to review the collective cladogram (Figure 4) as part of their reflective process.

During the second iteration, two of four sessions used paper handouts instead of tablet computers due to technology issues, and one session was prevented from receiving the intervention at all.

Iteration 2: Outcomes

Pre/Post-test

For iteration 2's assessment, we asked students the following questions:

1. What is evolution? How do you think it works?
2. How might evolution shape a species of red fox after 500 generations (approximately 1000 years)? Elaborate on the ideas from your previous answer. Feel free to speculate on the conditions surrounding their evolution.

The items were scored using a KI scale, designed to reveal deep conceptual understanding of evolution, which found that the mean post-test score (M=3.40, SD=0.92) was significantly greater than the mean pre-test score (M=2.58, SD=0.75), $t(24) = 3.91$, $p < 0.002$. An inter-rater reliability was performed on 17% of data, where 75% agreement (Kappa = 0.49, $p < 0.005$) was achieved at first pass. Once discrepancies were resolved 100% agreement was reached and the rubric revised to account for the differences.

Curriculum and Activity Artifacts

Prior to their participation in the collective inquiry activity, we found that of the 490 sections in the field guide, 84% were complete, however the wiki page documenting the rainforest at various time periods was only slightly more than half complete (58%). For the session that used tablet computers, we were able to examine how well students performed with the help of a guide when asked: *Which of the following is most likely your organism's ancestor?* Of 335 answers collected by the tablets, 84% were accurate. Similar to our results in iteration 1, there was an upward trend of accuracy compared to time period. However in iteration 2, over 90% accuracy was reached by "100 million years ago" whereas in iteration 1, over 90% accuracy was only reached at two time periods (25 and 2 million years ago; Figure 2). We also examined how student groups made comparisons between time periods, which included the following prompt: "As a team, you will compare the environment between 200 & 150 million years ago."

1. Discuss the following with your group members and record your answers below.
2. What are the major differences between the two time periods?
3. What species appeared in this time period that wasn't there before? Consider climate, habitat, animals, and plants.
4. What evolutionary processes might have occurred during this time period? How were these processes related to the climate, habitats or other species at the time?

In the sessions that were given paper handouts, only 67% of the comparisons were complete. While in the sessions utilizing tablets, all of the assigned comparisons were made. Student responses were scored using a KI scale, which found the mean KI score in the tablet sessions (M=3.33, SD=0.60) higher than that from the paper sessions (M=1.9, SD=0.69).

Due to time and technological constraints, students were unable to complete the final step discussing and posting ideas about evolutionary processes. This was remedied by asking students to answer an analogous question as a post-activity assignment: *Choose an organism from the Borneo ecosystem and discuss the evolutionary forces you think are at play thorough 200 millions years?* Responses were coded using the KI scale, which found that students who participated in the second iteration achieved a mean KI score of 3.7 (M= 3.74, SD=1.14). Over 50% of the notes were coded either as Basic or Complex (Figure 3).

Discussion

Disciplinary Content Knowledge

Pre/post assessments indicate that overall curriculum engaged students and helped them to learn within the evolutionary biology domain. Although the effects may be attributed to instruction of the entire unit, analysis of in-activity artifacts, especially those asking students to think deeply about underlying mechanisms of evolution and what constitutes as evidence for evolutionary processes, served as confirmation of our design's impact on disciplinary content knowledge. It is interesting to note that knowledge integration scores for students' explanation of evolutionary processes in iteration 2 were significantly higher than those in iteration 1. This may be due in part to the activity design or to their being completed after the activity, with more time for reflection. This could also be due to the phrasing of the question, which included more scaffolding in asking students to review the cladogram and choose about a specific species to reflect upon.

Engagement with Immersive Media

Reviewing students' participation and how well they paid attention to the media, observation accuracy rates suggest that students were paying attention to the content on the walls and those on the tablet. In the session utilizing paper handouts (in iteration 2), completion rate for making comparisons between time periods (e.g., 200 vs. 150 million years ago) were lower than the same question being answered on tablet devices, however

students in both session types were observed to be engaged with the content of immersive walls. One reason might be that on tablet devices, students did not have an option of not answering the question before moving forward in the activity. Another reason might be that the students in the paper group were asked to make comparisons between all time periods while the students in the tablet group were only asked to make two comparisons. It is also interesting to note that answers tended to be more complete with significantly higher KI scores (M=3.33, SD=0.60) than those in the paper group (M=1.90, SD=0.69), $p < 0.05$. Video analysis currently underway will serve to further elucidate how students engaged with the immersive media and how they influence student interactions that lead to productive knowledge co-construction.

Collaboration and Collective Progress

Structured Community Progress and Dependencies

We were able to structure community progress and student's dependency on their peers' work in two ways: 1) species specializations and teams with different specialists, and 2) giving students distinct roles during collaborative activities

1) A key design feature of the collective inquiry activity is the use of specializations (e.g., birds, primates) kept by the students throughout the curriculum. It offered students the opportunity to have an authoritative voice during collaborative discussions (since only they would have enlightened information about their specialty species). This was carefully designed into collaborative steps of the activity. It also allowed many students to work towards a shared artifact for further exploration. Over 50 students collectively completed a field guide in iteration 2. Several sessions of ten to fourteen students worked together to create a map of the evolutionary lineages of the species in the simulated rainforest ecosystem, and in both cases individual students had their own piece of the larger puzzle that they were responsible for. However, in order for them to really feel like "experts" in their specialization, students needed to engage in pre-activity assignments. Engagement was not enforced nor was effort graded, which might have led to lack of external motivation to complete tasks.

2) Another way we designed interdependency is to specify task-based roles within the activity (as opposed to content expertise). In iteration 2, when students collaborated to answer explanation questions, such as comparing time periods, they were asked to designate a specific person to act as the scribe of the group (for a particular question), as well time period specialists (i.e., responsible for looking up resources about the rainforest at each specific time). Another example of roles being designated was the use of guides in iteration 2.

Students Responding Meaningfully to Ideas of their Peers, in Real-Time

A more meaningful, but difficult, interdependency is to encourage students to respond meaningfully to ideas of their peers. This can be designed into the activity, to the extent that certain pieces of information may to be discovered by different people participating such that only by collaborating will they see the big picture. Relevant pedagogical information can be carefully scripted to emerge at the right time. In EvoRoom, different evolutionary mechanisms were designed for different species specialists to discover. For example, those specializing in plants and insects should become aware of the co-evolutionary relationship between flowering plants and pollen spreading insects, leading them to discover the symbiotic relationship between the fig tree and the fig wasp. This type of dependency was most difficult to achieve, since the revelation of one insight could rely on an emergent artifact that may or may not depict the correct information. The collective cladogram was one such artifact. The cladogram in both iterations showed conflicting information particularly in the earlier time periods. The facilitating teacher had to think swiftly on her feet when reading and trying to launch relevant discussions from emergent artifacts at the same time. However, our teacher demonstrated in both iterations that these emergent, aggregative artifacts can be a powerful teaching tool. Discrepancies in an emergent artifact led to interesting discussion. In one session, the mistakes in the earlier times of the cladogram prompted students to think about what constitute as evidence for evolution (e.g., fossil records), and that which records are more or less likely to survive, and what led to our current understanding of evolution.

Conclusion

Overall, we found that EvoRoom enabled students to think deeply about evolutionary biology mechanisms. With respect to the broader curriculum, assigning expertise (e.g., species) was an important aspect of the design but recognized that not everyone “bought-in” to their assigned roles, resulting in different levels of effort put into their work within the broader curriculum. This led to a varied range of expertise during collective inquiry activities (i.e., one “primate expert” having more insight than another), which would impact their group or sessions’ collective success. In both iterations, we found that students paid careful attention to the immersive media as well as other media components (tablet software, visualizations) in the immersive simulation. Much of the within-activity artifacts that demonstrated student reflection about evolutionary biology resulted from interactions with immersive walls as well as with discussion with students’ group members. However, this was as much a feature of our pedagogical design as the media’s influence. Further audio and video analysis is now

underway to measure students' thinking behind their written explanations. During the collective inquiry activity, visualizations that aggregated students' collective observations served as an important tool as a shared artifact to encourage deeper discussions and drive further reflection. From the early results, successful design features of our immersive simulation include: visualizing the rainforest, displaying the collective knowledge (i.e., in the form of a cladogram), and incorporating collaboration to the core of the activity.

References

Anderson, D., Fisher, K., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*(10), 952–978.

Birchfield, D., & Megowan-Romanowicz, C. (2009). Earth science learning in SMALLab: A design experiment for mixed reality. *Journal of Computer-Supported Collaborative Learning*, *4*, 403–421.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Brown, A. L. & Campione, J. C. (1990). Communities of learning and thinking, or a context by any other name. *Human Development, 21*, 108-125.

Colella, V. (2000). Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Sciences*, *9*(4), 471–500.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, *323*, 66–69.

Demastes, S., Good, R., & Peebles, P. (1995). Students' conceptual ecologies and the process of conceptual change in evolution. *Science Education*, *79*(6), 637–666.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., & Kirk, D. (2004). Savannah: mobile gaming and learning? *Journal of Computer Assisted Learning*, *20*, 399–409.

Horwitz, P., Schwartz, J., Kindfield, A., Yessis, L. M., Hickey, D., Heidenberg, A., & Wolfe, E. W. (1998). Implementation and evaluation of the GenScope™ learning environment: Issues, solutions, and results. Proceedings of the 3rd International Conference of the Learning Sciences, Charlottesville, VA.

Linn, M. C., & Eylon, B.-S. (2011). Science Learning and Instruction. *Routledge*. New York.

Lui, M., & Slotta, J. D. (2013). Immersive simulations for smart classrooms: exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education*, 1–24. doi:10.1080/1475939X.2013.838452

Mayr, E. (2001). *What Evolution is*. New York: BasicBooks.

Moher, T. (2006). Embedded Phenomena: Supporting science learning with classroom-sized distributed simulations (pp. 691–700). Presented at the CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, New York: ACM Press.

NSF Taskforce for Cyberlearning. (2008). *Fostering Learning in the Networked World: The Cyberlearning Opportunity and Challenge*. *Report of the NSF Task Force on Cyberlearning* (pp. 1–64). Arlington, VA: National Science Foundation. Retrieved from http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf

Peters, V. L., & Slotta, J. D. (2010). Scaffolding knowledge communities in the classroom: New opportunities in the Web 2.0 era. In M. J. Jacobson & P. Reimann (Eds.), *Designs for Learning Environments of the Future*. Secaucus, NJ: Springer.

Price, S., & Rogers, Y. (2004). Let's get physical: the learning benefits of interacting in digitally augmented physical spaces. *Computers & Education*, *43*, 137–151.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and Instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Erlbaum.

Rogers, Y., & Muller, H. (2006). A framework for designing sensor-based interactions to promote exploration and reflection in play. *International Journal of Human-Computer Studies*, *64*(1), 1–14.

Slotta, J. D. (2010). Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In K. Makitalo-Siegl, F. Kaplan, J. Zottmann, & F. Fischer (Eds.), *Classroom of the Future: Orchestrating collaborative spaces* (pp. 215–242). Rotterdam: Sense.

Slotta, J. D., & Chi, M. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, *24*(2), 261–289.

Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom*. Teachers College Press.

Slotta, J. D., & Najafi, H. (2012). Supporting collaborative knowledge construction with Web 2.0 technologies. In *Emerging Technologies for the Classroom* (pp. 93–112). New York, NY: Springer New York.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19. doi:10.1023/A:1009421303064

A Study of Subjective Emotions, Self-Regulatory Processes, and Learning Gains: Are Pedagogical Agents Effective in Fostering Learning?

Nicholas Mudrick, North Carolina State University, Raleigh, North Carolina, nvmudric@ncsu.edu
Roger Azevedo, North Carolina State University, Raleigh, North Carolina, razeved@ncsu.edu
Michelle Taub, North Carolina State University, Raleigh, North Carolina, mtaub@ncsu.edu
Reza Feyzi Begnagh, McGill University, Montreal, QC, reza.feyzibegnagh@mail.mcgill.ca
François Bouchet, Université Pierre et Marie Curie – LIP6, Paris, France, francois.bouchet@lip6.fr

Abstract: Though some research has focused on agent-direct affective processes, none has examined its impact on multi-agent learning environments and on the detection, modeling and fostering of self-regulated learning processes. 38 participants interacted with MetaTutor, an intelligent, multi-agent hypermedia-learning environment, to learn about the human circulatory system. The log files, containing information about their overall performance, and self-report measures, assessing emotions and impressions towards agents obtained from their interactions with MetaTutor were used to assess the relationship between subjective agent-directed emotions, SRL processes and overall learning gains. Results indicate that agent-directed emotions were not significantly related to SRL strategy use, negative agent-directed emotions were significantly related to negative learning gains and negative agent-directed emotions for two specific agents (representative of two SRL pillars) were related to negative learning gains. Implications for the design of multi-agent systems and the role of emotions during human-agent interactions and their relation to learning are discussed.

Introduction

While some researchers have focused on agent-directed affective processes, none have examined the impact that multi-agent systems have on the detection, tracking, modeling and ultimately fostering of different self-regulated learning processes. Within the literature, there is an implicit assumption that the use of pedagogical agents (PAs) is effective for learning and beneficial in fostering cognitive, affective, metacognitive and motivational (CAMM) processes (Azevedo et al., 2013). However, none have asked the fundamental question of whether or not PAs actually do facilitate the use of self-regulated learning (SRL) processes and impact overall learning about complex topics. Because of the paucity of current research regarding this topic, this study aims to address these issues through an examination of the associations between subjective emotions towards agents, SRL processes, and overall learning gains.

A recent review of the literature on PAs (Veletsianos & Russell, 2014) has revealed that research on PAs has been predominantly couched within socio-cognitive and cognitive-load theories, following the paradigm of computers as social actors (Kim & Baylor, 2006). A prevalent assertion within this paradigm is the view that PAs are able to aid learning with adaptive, supportive scaffolding and guidance of cognitive and metacognitive skills (Biswas, Leelawong, Shwartz, Vye & the Teachable Agents Group at Vanderbilt, 2005). Researchers within this area purport the possibility that agents can provide adaptive, individualized scaffolding and feedback to students and the basis for this idea rests in the success of one-to-one human-human tutoring (VanLehn, 2011). Another claim noted in the literature is the perceived ability of agents being able to emulate human behavior through realistic simulations (Sklar & Richards, 2010; Veletsianos & Russell, 2014). The employment of this strategy intends to facilitate natural communication between the PAs and the learners, and some research evidence indicates that when agents are highly realistic, an overall emotional connection is increased (Gulz, 2005; Woo, 2008). Furthermore, research has indicated that a sense of social interaction with the agents can help with collaboration in social learning (Gulz, 2005; Kim & Baylor, 2006; Sklar & Richards, 2010; Woo, 2008). Some studies indicate that engaging in unconstrained interaction with agents, learners treat the agents as conversational partners (Graesser & McNamara, 2010) while other research indicates that they could also lead to frustration and disappointment (De Angeli & Brahnam, 2008).

Contemporary research also operates on the expectation that the presence of agents increases students' motivation (Kim & Baylor, 2006; Kim & Wei, 2011; Kramer & Bente, 2010; Lusk & Atkinson, 2007). Agents are presumed to impact learners' interest, attention, and provide students a medium between humans and computers through the provision of motivational and affective instructions (Baylor, 2011; Lusk & Atkinson, 2007). Some have situated their examination into these processes within the persona effect, which claims learners perceive their learning experience positively as a result of understanding computers to be social actors (Choi & Clark, 2006).

A universal assumption within the research exists that PAs improve learning and performance through the combination of the aforementioned factors. Agents can provide various affordances that lead to a deeper understanding of the presented material (Veletsianos & Russel, 2014). Most importantly, empirical research "has shown that simply adding pedagogical agents in a digital environment does not lead to better learning outcomes," (Veletsianos & Russel, 2014, p. 764). Much of the literature has shown no significant differences between the presence and absence of agents on learning outcomes within an intelligent tutoring system (ITS) (Choi & Clark, 2006; Louwerse, Graesser, Lu, & Mitchell, 2005). While research that has found significance of agent presence has been predominantly focused on the design of the agents themselves (Jackson & Graesser, 2007; Veletsianos, 2007, 2010). All of the claims and assumptions about the effects of agents within ITSs are predicated on the students' interpretations of the agents themselves. However, the effects of these claims have yet to be addressed in a comprehensive manner.

We have begun to address some of these issues using an intelligent, hypermedia multi-agent learning environment called MetaTutor (Azevedo et al., 2009, 2010, 2012, 2013) that attempts to facilitate SRL in students as they learn about complex science topics. SRL rests on the assumption that successful learning is based in a learner's ability to accurately monitor and regulate their cognitive, affective, metacognitive, and motivational processes during learning (Zimmerman & Schunk, 2011). As such, it is a fundamental education construct that has been shown to be effective in students' ability to learn and study about various concepts (Azevedo, 2007; Winne & Perry, 2000; Zimmerman & Schunk, 2011). Research has shown that when students engage in SRL in the context of learning in computer based learning environments (CBLEs), they achieve high learning gains (Azevedo et al., 2013; Greene & Azevedo, 2010).

The impact of agent-directed self-regulation on learning gains with the MetaTutor environment has also been discussed. Within the learning environment, there are four PAs present who act to assist the student in learning through the provision of appropriate scaffolding for each participant. Each of the three agents specializes in one of the three pillars of self-regulated learning; Sam the Strategizer facilitates the use of learning strategies, Pam the Planner supports setting sub-goals and planning learning, and Mary the Monitor assists the learner monitor their overall learning. An additional agent, Gavin the Guide, functions to guide the participants through the learning environment as well as to prompt the self-report measures meant to assess the CAMM states of each participant. Several studies with the MetaTutor have revealed that these agents have a somewhat beneficial influence on learning and self-regulated learning processes (Azevedo et al., 2012, 2013).

However, we also have not examined the extent of the impact of PAs on SRL and learning gains, especially within the context of affect and SRL processes. It is of fundamental importance to look at the students' perceptions of the agents in relation to their performance. Their affective connotation towards the agents is a potential mediator of the agents' effectiveness in the promotion of self-regulated learning processes and overall learning. As such, the present study assesses the following research questions; (1) Are students' feelings towards agents associated with the frequency of their use of SRL processes; (2) What is the extent of the association between feelings towards the agent and the students' learning gains; (3) Are feelings towards the specific agents of MetaTutor (who are representative of specific SRL processes and the environment itself) related to learning gains?

Methods

Participants

A subset of 38 (68% female) undergraduate participants was sampled from a large multi-site research project. The sample included undergraduate students from two Canadian universities in Montreal, Canada, and one from Chicago, USA. The participants' ages ranged from 18 to 31 years old (M = 21.13, SD = 2.84) and were paid up to $40 dollars for completing the study.

Research Design

The data used in this paper comes from a larger database that involved an experimental study that examined the effectiveness of pedagogical agent facilitation of SRL in STEM education. There were two conditions in this study; a *control condition* and a *prompt and feedback condition*. In the control condition, participants were free to navigate the system, read content pages, and deploy SRL processes without any scaffolding or feedback from any of the pedagogical agents. In the prompt and feedback condition, participants were provided with prompts for use of SRL processes, scaffolding from the PAs, and received adaptive feedback and content from the PAs. A t-test was conducted to determine if there existed a significant difference between conditions on proportional learning gain. Results indicated the participants in the feedback condition (M = 31.09, SD = 30.14) did not significantly differ from those in the control condition (M = 40.24, SD = 35.92) in terms of proportional learning gain ($t(78) = 1.22$, $p = .225$). Because of this lack of statistical significant difference, only the members of the prompt and feedback condition were examined, as this was where the participants had the most interactions with the agents.

MetaTutor: An intelligent, Hypermedia Multi-Agent System

MetaTutor is an intelligent, multi-agent, hypermedia learning environment, which engages students in learning about a complex science topic, the human circulatory system (Azevedo et al., 2012, 2013). In MetaTutor, there are 38 pages of text and diagrams, all of which address different topics pertaining to the circulatory system. MetaTutor allows for the collection of a wide array of data, including log-file, eye-tracking, think-aloud data, electro-dermal activity (EDA), screen recordings of learner-system interactions, and facial expressions of participants' emotions. These multi-channel data were collected from students while they navigated the system and learned about the circulatory system. This aim of this analysis is to discuss the impact of agent-directed emotions and on self-regulated learning and overall learning gains.



Figure 1. The four pedagogical agents in MetaTutor



Figure 2. A Screenshot of MetaTutor

Experimental Set Up

Students were asked to participate in two sessions (the first, for one hour and the second, for three hours total) and were required to complete both sessions within three days of each other. During the first session, participants completed a consent form and were given an explanation of the study. Then participants began interacting with the environment, but only completing a series of self-report questionnaires that measured demographic information and their emotions. Participants then completed a 25-item multiple choice pre-test to assess prior knowledge of the human circulatory system and were paid $5 for completing this session.

In the second session, participants began interacting with the environment by first creating two out of the seven pre-determined sub-goals based on different aspects of the circulatory systems, which the pedagogical agent, Pam the Planner was programmed to recognize and lead the participants to set. Once the two sub-goals were established, Gavin the Guide presented multiple videos, which introduced the system, including all of the interface elements and how to engage in self-regulated learning strategies. Pam the Planner was presented again and prompted students to provide anything they knew about the given sub-goal they were working on. Finally, students began to learn by using the system freely and were able to use self-regulated learning strategies (e.g., taking notes, summarizing, etc.) and metacognitive monitoring and judgments (e.g., monitor their emerging understanding of the content, coordinate informational sources, etc.) at any point during the session by selecting which strategy they wished to from the SRL palette presented on the right of the screen. During learning with MetaTutor, a variety of different multi-channel data were collected, which included log-files, concurrent think-aloud protocols, EDA, facial expressions, eye-tracking, and audio-recordings, about each participants' self-regulated learning. Furthermore, the participants were also asked to provide responses to self-report measures like the Agent Persona Inventory (API), which was presented at the end of the session after the post-test, and the Emotions, Interests, and Values Questionnaire (EIV) that was administered every 14 minutes until a total of 5 EIVs were completed. At the end of the learning session, the participants were required to take a 25-item multi-choice post-test over the content presented during the session. The participant was then paid up to $40 for their completion of the study.

Data Sources: Learning Outcomes, Self-Report Measures, and SRL Behaviors

Two equivalent 25-item multiple choice pretests and posttests developed by Azevedo and colleagues (Azevedo et al., 2010) were used to assess participants' learning during the one-hour learning session with MetaTutor. A sample item was: "What is the effect of the clotting process?" Each item was scored as 0 or 1 depending on the accuracy of the answer (range for pretest and posttest was 0-25, respectively). Proportional learning gains were calculated using Witherspoon, D'Mello, and Azevedo's (2008) formula determine the amount of learning from pretest to posttest.

Two self-report measures were administered during the second session with MetaTutor, including the EIV and the API. The EIV consists of 20 questions that measure students' emotions, interests and values at a given time and is based Pekrun's (2011) model of academic achievement emotions. The responses are set on a Likert scale ranging from 1-5, where 1 is strongly disagree and 5 is strongly agree. A sample item was "Right now I feel happiness". The range of scores for the EIV is 20-100.

The objective of the API is to get the student to think about his or her interaction as a whole with each of the four agents. There are 22 questions that assess emotional states, with the 23rd a free response question asking what the agents could do differently to help. The student is asked to indicate the frequency with which each agent made them feel the state expressed in the question on a Likert scale of 1 to 5, where 1 indicated never and 5 indicated always. A sample item was "Sam made me feel angry", and this was repeated across the other three agents. The range of scores for the API is 88-440.

Data about learners' frequency of use of SRL processes during learning were extracted from the log-files, which captured the students' interactions with the MetaTutor environment. The extracted data consisted of the SRL strategies students engaged in, as well as their frequency of occurrence throughout the entire learning session. This was based on students' clicking on an SRL palette that was part of the system's interface. A sample button on the SRL palette was "Evaluate how well I already know this content".

Data Analysis

For this study, the data from API and EIV measures were combined into three categories, positive emotions, negative emotions and neutral. For the EIV, the positive emotions category included questions that asked "Right now I feel": happy, enjoyment, hope, pride, curiosity, eureka and surprise. The negative emotions category included a combination of "Right now I feel": anger, frustration, anxiety, fear, shame, hopelessness, boredom, contempt, disgust, confusion and sadness. The neutral category was the data collected from the question "Right now I feel neutral." Because the EIV was prompted a total of 5 times throughout the learning session, ratings data from each EIV administration for each respective student was combined to create a composite measure for overall affect when engaging in the MetaTutor learning environment.

The data collected from the API were combined in a similar manner to the EIV. Participants were asked, for each agent, how the agents made the participants feel. For example, "Sam made me feel happy." A positive emotions score for each agent was then combined, for the same emotions as combined for the EIV. Then, the negative emotions composite for each agent was created. The neutral composite was data from the measure assessing the extent to which all four agents "made me feel neutral." Finally, a compound measure was created for each emotion category. The positive emotions from each agent were combined into a total positive emotions score. The same process was done for the negative and neutral emotion categories.

Finally, the frequencies of cognitive and metacognitive processes for each participant were created into a total SRL category, which was a combination of measures that assessed the frequencies of judgments of learning, content evaluations, feelings of knowing, monitoring progress towards goals, prior knowledge activation, planning, taking notes, summarizing and inferring.

Results

Research Question 1: Are Students' Feelings towards Agents Associated with the Frequency of Their Use of SRL Processes?

Correlations were conducted to measure the association between proportional learning gains and the SRL composite scores. Table 1 demonstrates that the total composite score for SRL (r = .10, p = .536) was not significantly associated with proportional learning gain. The results tend to reveal that proportional learning and SRL strategy frequency were not significantly associated within this analysis. To further explain this association, the relation between the SRL composite and agent-directed emotions was then examined. Results indicated that positive agent-directed emotions (r = .29, p = .150), negative agent-directed emotions (r = .18, p = .377) and the neutral agent-directed emotion (r = .27, p = .096) were not significantly related to the SRL composite scores (see Table 2). Overall, this suggests that learners' agent-directed emotions were not significantly related to SRL frequency.

Table 1: Correlations between frequencies of SRL strategies and proportional learning gain

	Proportional Learning Gain	SRL Frequencies
Proportional Learning Gain	---	.104
SRL Frequencies	.104	---

Table 2: Correlations between frequencies of SRL processes and agent-directed emotions

	SRL Frequencies	Agent-Directed Positive Emotions	Agent-Directed Negative Emotions	Agent-Directed Neutral Emotion
SRL Frequencies	---			
Agent-Directed Positive Emotions	.238	---		
Agent-Directed Negative Emotions	.147	-.025	---	-.221
Agent-Directed Neutral Emotion	.274	.012	-.221	---

Research Question 2: What is the Extent of the Association between Feelings towards the Agent and the Students' Learning Gains?

Then, associations between agent-directed emotions and proportional learning gain were examined (see Table 3). The correlations revealed a significant, negative relationship between proportional learning gain and negative agent-directed emotions (r = -.38, p = .023), while the associations between positive and neutral agent-directed emotions (r = -.08, p = .624; r = .07, p = .640) remained non-significant. The more negative emotions learners experienced directed towards the agents, the worse their overall performance from pretest to posttest.

Table 3: Correlations between agent-directed emotions and proportional learning gain

	Proportional Learning Gain	Agent-Directed Positive Emotions	Agent-Directed Negative Emotions	Agent-Directed Neutral Emotion
Proportional Learning Gain	---			
Agent-Directed Positive Emotions	-.082	---		
Agent-Directed Negative Emotions	-.367*	-.025	---	
Agent-Directed Neutral Emotion	.078	.012	-.221	---

* p < .05 level.

Research Question 3: Are Feelings towards the Specific Agents of MetaTutor (Who Are Representative of Specific SRL Processes and the Environment Itself) Related to Learning Gains?

Because the relationship between negative agent-directed emotions was statistically significant, specific agent-directed negative emotions were then examined to determine whether or not any specific agent contributed to the negative relationship found. Within the four agents, the emotions towards two were found to be significantly and negatively associated with proportional learning gains. More specifically, results revealed negative directed emotions were found to be significant and negatively related to proportional learning gain for the male agents, Gavin (r = -.34, p = .036) and Sam (r = -.38, p = .018). While negative directed emotions towards the female agents, Pam and Mary (r = -.24, p = .15; r = -.27, p = .12) were not significantly related to proportional learning gains. These results suggest that negative emotions directed towards Gavin and Sam were associated with negative proportional learning gains.

Table 4: Correlations between specific agents directed emotions and proportional learning gain

	Proportional Learning Gain	Sam-Directed Negative Emotions	Pam-Directed Negative Emotions	Gavin-Directed Negative Emotions	Mary-Directed Negative Emotions
Proportional Learning Gain	---				
Sam-Directed Negative Emotions	-.383*	---			
Pam-Directed Negative Emotions	-.240	.721**	---		
Gavin-Directed Negative	-.342*	.623**	.597**	---	.

Emotions					
Mary-Directed Negative Emotions	-.256	.519**	.526*	.593**	---

** $p < .01$ level.
* $p < .05$ level.

Discussion

This study examined the often-neglected area of PAs, the emotions they induce in learners and their impact on SRL strategy use during complex learning with intelligent, multi-agent adaptive hypermedia systems. The results obtained from this study illustrate the potential impact that subjective feelings towards PAs have on students' learning gains within a multi-agent, adaptive hypermedia-learning environment. The analyses revealed that students' agent-directed emotions, at least negative ones, are associated with performance of SRL strategies and learning gains on macro, all agents considered, and micro, accounting for specific agents, levels.

The initial research question assessed the extent of the relationship between students' subjective feelings towards the PAs of the MetaTutor learning environment and SRL strategy frequency throughout their interaction within the environment. The correlation between SRL frequency and proportional learning gain indicated that, within this analysis, the two were unrelated. An examination was then conducted to assess whether an association between agent-directed emotions and SRL strategy frequency existed. No significant relationship was obtained. This result has tremendous implications for the incorporation of agents within CBLEs. One of the fundamental assumptions behind the inclusion of agents within CBLEs is that the PAs are helpful in fostering student SRL processes. However, results obtained from this analysis indicate this assumption more ambiguous than previously considered. This finding is bothersome since pedagogical agents are designed and embedded in technology systems based on the fundamental premise that learners' positive feelings towards them should be engaging, which should have led to a significant correlation between their feeling and the use of SRL processes and learning gains.

A second correlation analysis was conducted to determine whether a relationship existed between the students' emotions towards the agents were related to their proportional learning gains. The analysis indicated a significant, negative association between proportional learning gains and agent-directed emotions; the more negative the students' emotions towards the agents, the worse their proportional learning gain. This is understandable considering the agents' purpose themselves. As each agent is responsible for specific SRL strategies (i.e., planning, metacognitive monitoring, learning strategies), and negative emotions towards them could possibly interfere with learners inclination and compliance to enact the SRL strategies they are trying to promote the student to employ. If the students ignore the agents' SRL strategy suggestions, their performance and overall learning would be negatively affected. Furthermore, the presence of the agents themselves could account for this negative relationship between learning gain and negative, agent-directed emotions. Students' negative subjective feelings towards the agents could reflect their disdain for the agents' presence themselves, and this disdain could possibly impact their overall performance both through SRL strategy use and learning gain.

The results from third research question indicated similar results from those presented above. Negative emotions directed towards specific agents were examined in relation to proportional learning gain to determine if feelings towards any specific agent could account for this effect. Results indicated that negative emotions directed towards two of the agents, Gavin the Guide and Sam the Strategizer were negatively related to proportional learning gains. Contextualizing this finding within the framework of the MetaTutor learning environment explains these associations. Gavin is responsible for the provision of the self-report measures throughout the learning session. His presence is abrupt and the self-report measures he presents the students to fill-out during the session may detract from their attention to the ongoing learning about the human circulatory system. As such, specific negative feelings towards Gavin himself could be indicative of the student perceiving his presence as ultimately distracting, potentially negatively impacting performance and learning gain. Specific negative feelings towards Sam the Strategizer could be interpreted in line with his intended purpose within MetaTutor as well. Sam instructs students to perform specific SRL strategies, such as taking notes and summarizing. Once prompted by Sam, the student is required to use the strategies immediately prior to continuing to read and inspect additional hypermedia content about the body system. Due to learners' potential inability to comprehend or enact the effective learning strategy, they can go through many iterations of summarization before being allowed to resume interacting with the hypermedia material. As such, negative emotions directed at Sam could reflect a level of frustration and anger towards Sam that could divert students' attention from the task at hand and spiral into a cycle of negative emotions that require emotion regulation in order to deal with Sam's expectation of learner compliance. Furthermore, the negative feelings could also be indicative of the student not wanting to perform these SRL strategies, which, by not performing would ultimately negatively impact their overall learning gain. It is of interest to note that only the male agents were of

significance within the relationship between subjective feelings towards agents and learning gain. This suggests the operation of a potential gender bias on students' perceptions towards the agents as a whole.

Overall, this study provides several questions for future research. Theoretically, this study calls for an amalgamation of current models of self-regulated learning (e.g., Winne & Hadwin, 1998; Schunk & Zimmerman, 2011) and emerging models of externally-regulated learning (e.g., Hadwin et al., 2011) to incorporate social, cognitive, metacognitive, and affective mechanisms to account for the complex human-artificial agents' interactions during complex learning with pedagogical agents. Understanding the complex interactions between multiple pedagogical agents and learners is key to designing effective agents that are not only sensitive to students learning needs but also monitor and regulate their own behaviors (e.g., Sam recognizes a student's frustration when it asks the learner to makes inference and therefore does not repeatedly ask the learner to make inferences) so as to not negatively impact students' ability to monitor and regulate their learning and overall performance.

Methodologically, this study raises several questions about the measurement of self-regulatory processes and analytical approaches used to study complex agent-learner interactions. First, this study calls for the temporal alignment and convergence of multi-channel data (Azevedo et al., 2013). For example, eye-tracking data could be employed to determine the extent to which engaging with one of the four PAs results in specific emotions (e.g., frustration vs. confusion vs. neutral) and measure a tri-state affective cluster for each agent. More specifically, micro-level analyses can be conducted to determine affective transition state before interaction with an agent, during an interaction with an agent, and following the interaction with the agent. This type of data is key to understand the dynamics of emotion generation and regulation and provide fine-grained data on the impact of agents on learners' emotions and learning. Similarly, EDA data could augment eye-tracking data to determine physiological-emotional correspondents while interacting with agents.

In terms of analytical approaches, this study used molar-level aggregation of learning outcomes and self-report measures. A major challenge for the interdisciplinary field of learning sciences remains the development and testing of more sophisticated analytical methods that can be used to describe the underlying phenomena (e.g., impact of agents' presence on the fluctuation of students' affective states during contextually-bound episodes of learning). In addition, the overreliance on self-report measures as the dominant measure of affective and motivational processes must be augmented with processes-oriented measures of emotions. Also, clusters across self-report scales or a trifurcation of positive, negative and neutral emotions could be explored in addition to specific learning centered emotions themselves.

The implications for these findings are staggering and have been left largely unexplained by contemporary research. Assessing whether PAs within CBLEs are effective in what they are designed to do, on a larger scale within different learning environments, has yet to be asked. Contemporary research rests on the assumption that PAs are effective in facilitating the use of SRL strategies, yet the results presented here are in direct conflict with that expectation. Furthermore, the inclusion of PAs is presumed to increase performance and overall learning, while the outcomes of this paper indicate a negative relationship. As these results suggest, including agents for the sake of including agents could potentially be deleterious to students' overall learning.

References

Azevedo, R. (2007). Understanding the complex nature of self-regulatory processes in learning with computer-based learning environments: an introduction. *Metacognition and Learning, 2*, 57-65.

Azevedo, R. (2009). Theoretical, methodological, and analytical challenges in the research on metacognition and self-regulation: A commentary. *Metacognition & Learning, 4*(1), 87–95.

Azevedo, R., Feyzi-Behnagh, R., Duffy, M., Harley, J., & Trevors, G. (2012). Metacognition and self-regulated learning in student-centered learning environments. In D. Jonassen and S. Land (Eds), *Theoretical foundations of learning environments*. NY: Routledge.

Azevedo, R., Harley, J., Trevors, G., Duffy, M., Feyzi-Behnagh, R., Bouchet, F., & Landis, R. S. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agents systems. In R. Azevedo & V. Aleven (Eds.). *International handbook of metacognition and learning technologies*. Amsterdam, The Netherlands: Springer.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. In M. S. Khine and I. M. Saleh (Eds.). *New science of learning: Computers, cognition, and collaboration in education*, 225-247. Heidelberg: Springer.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development, 59*(2), 291-300.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at Vanderbilt. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, *19*, 363-392.

Choi, S., & Clark, R. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of Educational Computing Research, 34*(4), 441-466.
De Angeli, A. & Branham, S. (2008), I hate you! Disinhibition with virtual partners. *Interacting with Computers, 20*(3), 302-310.
Graesser, A., & MacNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45*(4), 234-244.
Greene, J., & Azevedo, R. (2010). Theoretical, conceptual, and methodological issues in the measurement of cognitive and metacognitive processes during learning with computer-based learning environments. *Educational Psychologist, 45*(4), 203-209.
Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.). *Handbook of self-regulation of learning and performance*, 65-84. NY: Routledge,
Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Pery, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36-48.
Gulz, A. (2005). Social enrichment by virtual characters-Differential benefits. *Journal of Computer Assisted Learning, 21*, 405-418.
Jackson, G. T. , & Graesser, A. C. (2007). Content matters: An investigation of feedback categories within an ITS. In R. Luckin, K. Koedinger, & J. Greer (Eds.), *Artificial intelligence in education: Building technology rich learning contexts that work*. Amsterdam: IOS Press.
Kim, Y., & Baylor, A. (2006). A socio-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development, 54*(6), 569-596.
Kim, Y., & Wei, Q. (2011). The impact of learner attributes and learner choice in an agent-based environment. *Computers in Education, 56,* 505-514.
Kramer, N. C., & Bente, G. (2010). Personalize e-learning: The social effects of pedagogical agents. *Educational Psychology Review, 22,* 71-87.
Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology, 19*(6), 693-704.
Lusk, M.,& Atkinson, R. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology, 21,* 747-764.
Sklar, E., & Richards, D. (2010). Agent-based systems for human learners. *The Knowledge Engineering Review, 25*(2), 111-135.
VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197-221.
Veletsianos, G. (2007). Cognitive and affective benefits of an animated pedagogical agent: Considering contextual relevance and aesthetics. *Journal of Educational Computing Research, 36*(4), 373-377.
Veletsianos, G. (2010). Contextually relevant pedagogical agents: Visual appearance, stereotypes, and first impressions and their impact on learning. *Computers in Education, 55*(2), 576-585.
Veletsianos, G., & Russell, S. R. (2014). Pedagogical agents. In J. M. Spencer, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 759-769). New Yotk, NY: Springer.
Woo, H. L. (2008). Designing multimedia learning environments using animated pedagogical agents: Factors and issues. *Journal Computer Assisted Learning, 25*, 203-218.
Winne, P. H. , & Hadwin, A.F. (1998). Studying as self-regulated learning. In D.J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Hillsdale, NJ: Erlbaum.
Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531-566). Orlando, FL: Academic Press.
Zimmerman, B. & Schunk, D. (Eds.) (2011). *Handbook of self-regulation of learning and performance*. New York: Routledge.
Witherspoon, A., Azevedo, R., & D'Mello, S. (2008). The dynamics of self-regulatory processes within self- and externally-regulated learning episodes. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems: Lecture Notes in Computer Science (LNCS 5091)* (pp. 260–269). Berlin: Springer

Acknowledgements

The research presented in this chapter has been supported by funding from the National Science Foundation (DRL 1008282), the Institute of Educational Sciences (R305A120186), the Social Sciences and Humanities Research Council of Canada (430–2011–0170 and 890–2012–0138), the Canadian Foundation for Innovation, and the Canada Research Chairs program awarded to the second author.

Design-Based Research Process: Problems, Phases, and Applications

Matthew W. Easterday, Daniel Rees Lewis and Elizabeth M. Gerber
Northwestern University, Evanston, IL
{easterday, daniel-rees, egerber}@northwestern.edu

Abstract. Since the first descriptions of design-based research (DBR), there have been continued calls to better define DBR and increase its rigor. Here we address four uncertainties about DBR: (a) the phases of the DBR process, (b) what distinguishes DBR from other forms of research, (c) what distinguishes DBR from design, and (d) the characteristics of DBR that make it effective for answering certain types of questions. We build on existing efforts by defining DBR as an iterative process of 6 phases: focus, understand, define, conceive, build, and test, in which other scientific processes are recursively nested. By better articulating the process of DBR, this definition helps us to better craft, improve, communicate, and teach design-based research.

Introduction

Although design has existed since the beginning of human history, its rise as an educational research methodology is relatively recent. Descriptions of design-based research (DBR) in education include: Brown (1992), special issues of Educational Researcher (Kelly 2003), and the International Journal of Learning Sciences (Barab & Squire, 2004), and several edited volumes (Kelly, Lesh, & Baek, 2008; Plomp & Nieveen, 2007; Van den Akker 1999; Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006). After several decades of work on DBR, some have concluded that: "as promising as the methodology is, much more effort … is needed to propel the type of education innovation that many of us feel is required" (Anderson & Shattuck, 2012). While it is difficult to evaluate an entire research methodology (McKenney & Reeves, 2013), proponents of DBR should take these criticisms as a friendly challenge to more rigorously define DBR (Hoadley 2004).

There is general agreement that DBR should generate effective educational interventions and useful theory (Van den Akker et al., 2006, Ch. 1). We consider DBR to cover a wide range of projects, such as Margolis and Fisher's ethnographic study of women in computer science that produced a theoretical model used to re-design a computer science department, increasing the percentage of women from 7% to 42% over 5 years (2003, p. 6). In our own work, this combination of design and research includes formative evaluations and controlled randomized experiments that result in design principles for educational games that allow us to increase both learning and interest.

DBR provides educational researchers with a process for *use-inspired basic research* (Stokes 1997; Schoenfeld 1999; Lester 2005) where researchers design and study interventions that solve practical problems in order to generate effective interventions and theory that is useful for guiding design. DBR is important because it recognizes that neither theory nor interventions alone are sufficient. The classical model of research and development, that is, basic research leading to applied research, leading to development, leading to products, does not work well (Stokes 1997). Alternatively, design, unguided by theory, is likely to be incremental and haphazard. Theory derives its purpose from application and application derives its power from theory. Our problem as DBR researchers is to devise a means of conducting DBR that reliably produces both theory and interventions.

Problems arising from the ill-definition of DBR

Unfortunately, there are many unresolved issues with DBR that arise because we lack a clear definition about what DBR is, how it is conducted, and what it produces. We describe four of these problems.

Problem 1: Uncertainty about the DBR Process

The first problem is the uncertainty about the phases of DBR--the process typically looks different depending on who conducts it. There seems to be no accepted precisely described DBR process at the level of specificity dedicated to other methodologies such as experiments or grounded theory.

Understanding the DBR process requires us to define the phases of DBR. A *phase* describes the goal of a set of *methods* within a design *process*; for example, surveys and interviews could be considered methods in a data collection phase of a research process. We need to understand the phases of design so that we can: make coherent decisions about which methods to apply and when; explain the high-level process of DBR to new researchers; effectively communicate DBR methodology in the concise form required for publication; and understand similarities and differences across different instantiations of DBR in a way that allows us to borrow

 © ISLS

methods and improve the DBR methodology. Understanding the phases of DBR allows us to better design and to better communicate.

The integrative learning design framework (ILDF) (Bannan 2007; Bannan-Ritland 2003) is perhaps the best attempt to define the phases of DBR. However, the four phases in ILDF: exploration, enactment, local impact evaluation and broader impact evaluation, blend distinct design goals. For example, the enactment phase includes *prototyping* and the evaluation phase includes *system refinement*--both of which have a similar goal of building an intervention, but which nevertheless appear in different phases. Furthermore, the local evaluation phase and broader impact evaluation phase conflate the phase goal (of evaluating) with iteration. That is, both small-scale evaluation and large-scale evaluation serve the goal of testing, they simply occur in earlier or later iterations. Conflating phase and iteration creates a problem when we imagine an intermediate evaluation between local and broader impact that cannot be fit into the framework. Finally, it is not clear "where in this framework might randomized field trials be appropriate" (Bannan-Ritland 2003, p. 24).

Some of the most popular design processes used by practitioners like *Instructional Systems Design* (Dick, Carey, & Carey, 2008) provide a clearly articulated process and methods for designing instruction but do not attempt to define the high level phases of design or how the process might be used for research. Other popular design frameworks such as ADDIE (Analysis, Design, Development, Implementation, Evaluation) provide an umbrella term but "no real or authentic meaning" (Molenda 2003, p. 36).

Problem 2: Uncertainty about How DBR differs from Other Forms of Research

DBR is typically imagined as a form of *qualitative* research useful for building theory, that is, for addressing the problem of meaning (Kelly 2004) or used in the context of discovery (Kelly 2006, p. 177) as opposed to verifying an existing theory. While qualitative, it is distinct not just from laboratory experiments but also from ethnography and large-scale trials (Collins, Joseph, & Bielaczyc, 2004). Others argue that DBR can be productively *interleaved* with quantitative methods, for example, as a mixed methods approach crossing the field and lab (Brown 1992, p. 152-154; Kelly 2006, p. 169-171), as a point on an interleaved continuum (Hoadley 2004), or as a methodology with an agnostic stance toward quantitative and qualitative perspectives (Bannan-Ritland 2003, p. 24). Other writings describe DBR as a way to *integrate* other research methods (Collins, Joseph, & Bielaczyc, 2004, p. 39) or disciplines (Buchanan 2001) and that "methods of development research are not necessarily different from those in other research approaches" (Van den Akker 1999, p. 9). These research methods are applied in a stage appropriate manner (Bannan-Ritland 2003; Kelly 2004, 2006, p. 177). Finally, there is disagreement amongst design research theorists (outside of education) about whether design is a **science** at all, with some arguing that it is a science focused on the nature of designed artifacts (Simon 1996), others arguing that such a science is impossible because designers address problems that are not generalizable (Buchanan 1992, p. 17).

Problem 3: Uncertainty about How DBR Differs from Design, or Why Design Is Not Research

DBR proponents seek to establish DBR as a distinct and valid form of research. However, in arguing for DBR, we often ignore how DBR differs (if at all) from design as practiced in industry. Other fields, such as human-computer interaction, struggle with similar questions (e.g., Zimmerman, Forlizzi, & Evenson, 2007). Researchers claim DBR differs from design because it is: (a) research driven, that is, it addresses research questions, references literature, produces theoretical claims, and seeks to generalize beyond a specific context; and (b) involves more systematic evaluation, including formative data collection, documentation and analysis, (Bannan 2007; Edelson 2002).

Bannan (2003) points out that these are not typical attributes of practitioner methodologies like ISD (Dick, Carey & Carey, 2008). Of course, designers in industry often use qualitative methods (e.g., Beyer & Holtzblatt, 1998); develop novel, generalizable interventions described in forms such as patents or software patterns; rigorously evaluate qualitative and quantitative data through user-testing labs (Thompson 2007) and large scale experiments such as Google's A/B testing (Christian 2012). It is not clear whether there is a clear separation between design and design research or whether the distinction is artificial, or somehow peculiar to the field of education.

Problem 4: Uncertainty about What Might Make DBR Effective (If It Is)

The lack of clarity about the nature of DBR makes it difficult to justify its effectiveness as a research methodology. DBR is only useful if it allows us to reliably produce useful interventions and effective theories, "better, faster, or cheaper" than other methodologies, or to do so at least in some contexts. Without a clear description of the DBR process, we cannot make a coherent argument about the tradeoffs between DBR and other methodologies.

To increase the rigor of DBR, we need to provide a formal definition of DBR. The 4 problems arise because we do not have a clear definition about how DBR is conducted, at least not at the level of specificity provided for other methodologies. In 1992, Brown called on the field to define DBR and a decade later special issues in Educational Researcher and IJLS set out to answer that call; two decades later, we still lack a clear definition. DBR remains what organizational behavior researchers call a *low paradigm* field (or practice), where there is little technical consensus about the research questions considered important, the guiding theoretical models and, most significantly for our purposes, research methods (Pfeffer 1993). Low paradigm fields have more difficulty acquiring funding (because funders can be less certain of results), have lower journal acceptance rates (because there are greater disagreements about quality), lower collaboration and more difficulty training graduate researchers--all ultimately resulting in lower accumulation of knowledge (Pfeffer 1993; and Herrington, McKenney, Reeves, & Oliver, 2007 on DBR doctoral training). Reasonable people might disagree about the paradigmatic status of the Learning Sciences, but the calls to better define the argumentative grammar (Kelly 2004) and rigor (Hoadley 2004) of DBR suggest that we can make DBR a higher paradigmatic practice. Dede puts it bluntly: "...neither policy makers nor practitioners want what the DBR community is selling right now. We appropriately don't match the narrow conceptions of science currently in vogue at the federal level, but have much internal standard-setting to accomplish before we can put forward a defensible alternative" (2004, p.14). Twenty years on from Brown and Collins, the benefits of increased methodological consensus warrant a renewed attempt to provide a formal definition of DBR.

A Formal Definition of the Design-Based Research Process

Here we present a definition of DBR as a process that integrates design and scientific methods to allow researchers to generate useful products and effective theory for solving individual and collective problems of education. This paper focuses on describing the DBR process as part of this definition.

Design Process

The design and DBR processes consist of 6 iterative phases in which designers: *focus* the problem, *understand* the problem, *define* goals, *conceive* the outline of a solution, *build* the solution, and *test* the solution (Figure 1).

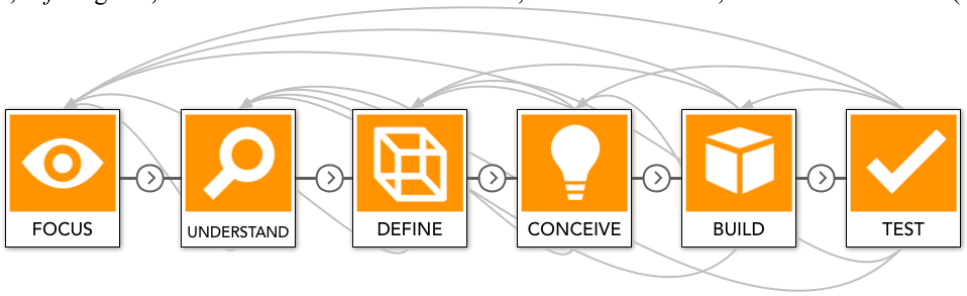


Figure 1. The design process consists of 6 iterative phases: focus, understand, define, conceive, build and test.

Focus

In the focus phase, designers bound the audience, topic, and scope of the project. The audience specifies whom the product serves, including learners and the other stakeholders affected, such as parents or the community. The team specifies who is designing the product and their reasons for participating. The topic specifies the general problem the product should address and how it arose. The scope specifies the constraints and the scale of the project. These issues are typically captured in a design brief.

Why: Focusing sets the direction of the project. A design is meant to achieve an intended goal and there can be no meaningful goal without some problem or opportunity to address. Focusing ensures that there is something worth designing and that the team has the expertise to succeed.

Understand

In the understand phase, designers study learners, domains, contexts and existing solutions. The understand phase investigates the problem through empirical methods and secondary sources, and synthesizes that knowledge into a form that can be easily used later in the process. Empirical methods include quick human-centered techniques such as observation, interviewing, surveys, data analytics, etc. Review of secondary sources focuses on: research that helps understand the problem such as models of learning and cultural contexts; analysis of current solutions to similar or related problems; and identification of design principles. The empirical data and research literature must be synthesized through methods such as identifying themes, building graphical models and creating learner personas.

Why: Typically the initial impetus for the project involves a situation in which existing solutions do not work or for which a novel solution is desirable--so designers must work to understand the nature and causes of the problem. Applicable secondary sources can be tremendously helpful in understanding the problem or

avoiding dead ends, but typically the problem arises in the first place because the root causes are unclear or because existing knowledge is insufficient to solve the problem. Furthermore, design requires detailed knowledge of user needs and context so empirical methods that can be employed quickly are almost always necessary to understand the problem.

Just as in science, discovering new features of the learning environment in the understand phase may be the core innovation of the design or theoretical contribution, such as building a better model of expertise or identifying the learning challenges in a particular domain. This includes ontological innovations, such as identifying Meta-Representational Competence (diSessa & Cobb, 2004) as a needed skill in a domain.

Define

In the define phase, designers set goals and assessments. Defining means converting an indeterminate problem, which has no solution, into a determinate problem that can be solved (Buchanan 1992). There are many ways to frame a problem. For example, suppose that the designer finds that: (a) the target learners are from immigrant communities, (b) their client wants to improve learners' performance on common core literacy and civic education standards, and (c) there are gaps in research literature about how to leverage learners' cultural resources. The problem could be defined as a question of "how might we engage students in debates about legal status?" or "how might we teach students to construct video documentaries about immigration policy?" or "how might we teach students to analyze the political values in English/Spanish-language youth media?" By completing the sentence "How might we...?" the designer selects a goal from the infinite and unknown number of goals that *could* be defined.

Why: A design focus, by definition, cannot be solved because there is no determinate (specific) goal provided--that is, there is nothing explicit to solve. It is up to the designers to define what that goal is, taking into account the goals important to the stakeholders and which can be productively solved. Only after the goal has been defined can a design be said to succeed or fail.

A novel problem definition can be the core innovation because it can lead to entirely new kinds of solutions.

Conceive

In the conceive phase, designers sketch a plan for the solution. Given a *definition* (even if implicit) the designer can plan a design intended to reach the goal. This involves imagining a solution and analyzing whether it will work. In this phase, the designer has not committed to implementing the design in a given medium, but rather creates a non-functional, symbolic or graphical representation that allows the designer to conceptually analyze the solution by determining the components of the design and how they might work together. Here, designers also develop theoretical products (diSessa & Cobb, 2004) such as design arguments (Van den Akker 1999), the underlying principles of the product, which may be of different levels of complexity (Buchanan, 2001), from communication, to artifacts, services, and systems (Penuel, Fishman, Haugan Cheng & Sabelli, 2011). The distinction between the conceive and build phase is between that of a conceptual plan constrained only by the designer's knowledge and that of a concrete prototype that is at least partially functional and constrained by a medium.

Why: Designers have a number of tools for planning, sketching, and modeling a design. These tools allow designers to test the design against their own knowledge and theory, to identify problems and improved solutions before committing to implementation in a particular medium, which can be difficult, costly, or time consuming.

Build

In the build phase, designers implement the solution. Once a design has been conceived, the designer can implement the design in a form that can be used. This implementation can be of lower or higher fidelity depending on the stage of the project and the question that the designer wants to test, which may be about a particular aspect of the educational intervention, or whether the educational intervention as conceived can achieve its goal.

Why: A design must be implemented to achieve a goal, and because a design is never completely finished, every implementation provides a prototype that can answer questions about whether the goal has been achieved.

Test

In the test phase, designers evaluate the efficacy of the solution. Iterative user-testing involves testing successive (often parallel) versions of the design at increasing levels of fidelity. Early testing of the plans produced in the conceive phase focuses on questions of relevance and consistency and then later on expected practicality, with expert reviews and walkthroughs. Later testing on prototypes constructed in the *build* phase

focus on questions of actual practicality and effectiveness using 1-1, small group, field trials and their variants (Tessmer 1993).

Testing often uses formative evaluation, which may not establish causality to the extent possible in controlled, randomized experiments, but which can quickly reject bad designs. This increases the likelihood of finding an effective design that can be verified later through summative evaluation. Some consider the boundary between formative and summative evaluation the point at which design research ends and the sciences of the artificial (Simon 1996), or in this case, rigorous evaluations testing strong causal claims of design principles, begins. We consider both valid forms of testing in DBR.

Why: Testing provides the designer with feedback about the success of the design and the validity of the theoretical propositions. It tells the designer whether the design has achieved its practical and theoretical goals.

Iteration

The design phases are not carried out in a linear sequence but rather iteratively. For example, in *building* an educational game, formative *testing* might reveal that the game is only attractive to boys, so one might return to *understand* how gender affects the likability of specific game features.

Rapid iteration is a tenet of modern human-centered design. It protects against the risks of designing interventions that are over-budget and behind schedule by quickly testing the designer's assumptions. Rather than design an entire intervention and discover only at the end that it does not work, iterative design argues for quickly building low fidelity prototypes, testing them, and re-designing--gradually evolving the intervention over time.

There is a delicate balance between planning, iteration and medium. When planning allows designers to avoid mistakes and the medium makes testing costly (e.g., building bridges), then there will be little iteration or at least a greater emphasis on lower-fidelity prototyping and modeling. However, if our ability to avoid bad designs through planning is limited and the medium makes the costs of testing low (e.g., web applications), then iteration is likely to be quick and frequent. Because education is a complex environment, our ability to predict the effect of an intervention is low. The cost of testing in education is probably relatively moderate--while the cost of implementing a lesson is low, the cost of testing may be greater depending on the type question/evaluation.

The DBR Process Includes Recursively Nested Research Processes

Scientific findings are also *products* created (or discovered) by a design process. For example, scientists may conduct an experiment in which they *focus* on a topic, *understand* the background literature, *define* a hypothesis, *conceive* of an experiment, *build* evidence by gathering and analyzing data, and finally *test* the validity of their findings, perhaps through peer-review. Qualitative research methodologies such as grounded theory follow a similar set of phases, except there the purpose is to *build theory* rather than verify a hypothesis.

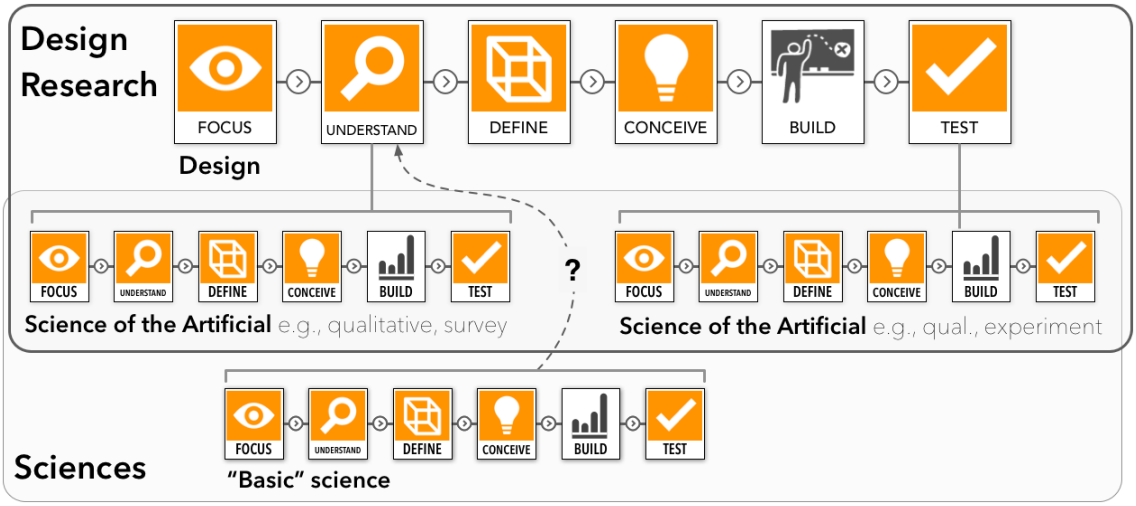


Figure 2. Scientific research methodologies (both qualitative and quantitative) follow a design process and produce products such as theories and models that can be incorporated into the design of another product such as an educational intervention.

Products that serve one purpose, such as verification of a hypothesis, can be used as components in the design of another product, such as an educational intervention (Figure 2). That means that in designing a learning environment, we might conduct other sub-design processes (such as a qualitative study or an

experiment) as part of the DBR process. For example, a DBR study of a journalism curriculum might conduct a qualitative study about learners' media practices in the understand phase, or a controlled randomized test of the curriculum in the test phase. In other words, design processes can be recursively nested within each other. This explains the shape-shifting nature of DBR--DBR looks like other forms of research because it incorporates these methodologies to do its work.

Stage Dependent Search

By understanding how design incorporates other scientific design processes, we can make a more compelling argument for why DBR can be an effective educational research methodology. Design research uses a stage-dependent search strategy (Bannan-Ritland 2003; Kelly 2004, 2006), in which designers choose different build and test methods depending on the stage of the design. In early stages of a project, such as when the problem context is poorly understood and there are few effective implementations, researchers are likely to produce unsuccessful designs, so they must choose a research and development strategy that allows them to quickly reject failures and understand the theoretical issues that must be addressed. So in the early stages of a project, researchers should focus on low-fidelity prototyping and collecting the minimal amount of data needed to quickly reject failure and identify potential successes. As researchers identify promising prototypes they can focus on theory building with qualitative methods to better understand the issues a design might address and the mechanism through which it affects learning. Once researchers have a plausible, well-grounded theory and an implementation with some evidence of success, they can conduct randomized controlled experiments to verify the efficacy of the theory and intervention. If researchers use randomized, controlled, experiments at the beginning stages of a complex design problem, they are likely to waste resources verifying a bad design. Likewise, if researchers never advance beyond theory building and radically novel designs, they are unlikely to provide strong evidence for the efficacy of an intervention or principle.

Resolving the Uncertainties

This formal definition of the DBR process resolves the uncertainties presented earlier.

Problem 1 resolution*: a clear definition of the phases of DBR.* The formal definition resolves the uncertainty about the phases of design in a way that allows us to better conduct DBR, train new researchers, improve DBR methodology, and communicate process within and outside the DBR community.

Problem 2 resolution: *DBR differs from other research in that it designs a product while using other methodologies as nested processes (sub phases) of design.* The formal definition shows how DBR differs (or rather does not differ) from other forms of research. DBR incorporates other scientific design processes into the design process for creating educational interventions in a recursive, nested manner.

Problem 3 resolution: *DBR differs from design practice in that it does not just produce an educational intervention but makes use of nested scientific processes to produce theory.* The formal definition also shows how DBR differs from "normal" design. By incorporating scientific processes, DBR produces theories connected to the literature and more rigorously tests interventions. Of course, there is no hard line separating the work of practitioners and researchers because practitioners use similar methods--the difference is one of degree and intent.

Problem 4 resolution: *DBR produces gains by deploying the appropriately nested scientific process at a given stage of development.* The formal definition shows how DBR efficiently develops theory by quickly identifying plausible interventions and constructs in early phases that are more rigorously verified in later stages.

Applying the Definition

A better understanding of the DBR process helps us to do better design research, train new researchers, improve DBR methodology, and communicate process within and outside the DBR community.

Better Design

Defining the DBR process helps us to better determine which methods to use and when. For example, when planning DBR projects, thinking about the *test* phase has prevented us from jumping to formal evaluation too early or dwelling in theory building too long. For ill-defined problems, we have used the phases to justify spending more effort applying methods from the understand phase. The phases also make clear when we have only implicitly defined the goals and design arguments for a project. DBR projects work under constraints of people, resources, and time, and the phases have allowed us to more deliberately deploy those resources.

Training New Researchers

There is a bewildering array of methods applicable in DBR projects and it is challenging for new researchers to make sense of these methods (Herrington et al., 2007). We use the DBR phases to explain how the design research process works at a high level, to help novices organize sets of research methods, and to explain the

meta-cognitive strategies we use to conduct design-based research. Just as design phases help researchers think precisely, they also serve as a tool to make design logic explicit to new DBR researchers.

Improving DBR Process

A clear definition of the DBR phases also helps us to improve the process. In struggling to consolidate learner data gathered in the *understand* phase, we have used human-centered design methods for synthesizing user data, such as personas. Or in rethinking curricula as services, we have applied *conceive* methods from service design such as journey maps, swimlanes and service blueprints. The phases allow DBR researchers to more easily borrow methods from other methodologies just as human-centered design has borrowed methods from methodologies such as ethnography. The DBR phases serve as a Rosetta Stone for translating and synthesizing design processes from other methodologies.

Likewise, we can use the phases as an analytical tool for judging design processes and potential contributions. For example, noticing that the ADDIE process does not clearly identify *focus* and *define* stages, or that the ILDF conflates phase and iteration. By identifying gaps, the design phases allow us to suggest new methods that can be applied to improve these processes. Furthermore, each phase identifies the locus of potential design research contributions when clearly defined.

Communicating Research Process

We have also used the phases to describe the choices made during a DBR project and why those were effective. In publishing research and grant applications, the phases more concisely communicate the past history or future plans of a DBR project. Unfortunately, the lack of shared vocabulary and conventional methodology creates a communication barrier, for example, in grant applications that require lengthy descriptions of planned cycles of design, iteration and testing.

Well-defined DBR phases allow us to explain the logic of DBR to other researchers. For example, quantitative psychologists may see the lack of inter-rater reliability in the early stages of a DBR project as a lack of rigor. Researchers from other disciplines will naturally judge DBR by the methodological standards of their own discipline. However, when DBR researches explain the methodological logic of shifting from an early focus on design concepts and theory building to a later focus on verification, we've found that those outside the discipline are often sympathetic to the aims of DBR. The problem is not that researchers from other disciplines are unaware of the methodological challenges of developing new interventions and theories (which DBR was developed to address), the problem is that other researchers will only accept DBR's alternative approach to addressing these methodological challenges when DBR researchers clearly and precisely articulate the rationale behind the DBR methodology.

Conclusion

We have defined DBR a**s** a process that integrates design and scientific methods to allow researchers to generate useful educational interventions and effective theory for solving individual and collective problems of education. This definition of the DBR process is neither "a way" nor "the way" to conduct DBR, rather, it describes the fundamental nature of all forms of DBR in order to help us better communicate and think about DBR. This definition is not just an academic exercise, but necessary to establish DBR as a high paradigm methodology, allowing us to better replicate the design process, to apply methods from other design methodologies, to better teach DBR to new design researchers, to acquire more resources, and ultimately to accumulate theory relevant to practice. By formally defining DBR, we establish its credibility as a legitimate methodology of educational research.

References

Anderson, T., & Shattuck, J. (2012). Design-Based research A decade of progress in education research? *Educational Researcher*, *41*(1), 16-25.

Bannan, B. (2007). The integrative learning design framework: An illustrated example from the domain of instructional technology. In T. Plomp & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 53-73). Netherlands Institute for Curriculum Development.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher*, *32*(1), 21-24.

Barab, S., & Squire, K. (2004). Introduction: Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1-14.

Beyer, & Holtzblatt. (1998). *Contextual design: Defining customer-centered systems.* San Francisco: Morgan Kaufmann.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141-178.

Buchanan, R. (1992). Wicked problems in design thinking. *Design Issues*, *8*(2), 5-21.

Buchanan, R. (2001). Design research and the new learning. *Design Issues*, *17*(4), 3-23.
Christian, B. (2012, April 25). The A/B test: Inside the technology that's changing the rules of business. *Wired.*
Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15-42.
Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; Disessa and Cobb; and Fishman, Marx, Blumenthal, Krajcik, and Soloway in the JLS special issue on design-based research. *The Journal of the Learning Sciences*, *13*(1), 105-114.
Dick, W., Carey, L., & Carey, J. O. (2008). *The systematic design of instruction* (7 ed.). New York: Pearson.
diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences*, *13*(1), 77-103.
Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences*, 11(1), 105-121.
Herrington, J., McKenney, S., Reeves, T., & Oliver, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. In C. Montgomerie & J. Seale (Eds.), *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2007 (pp. 4089-4097). Chesapeake, VA: Association for the Advancement of Computing in Education.
Hoadley, C. M. (2004). Methodological alignment in design-based research. *Educational Psychologist*, *39*(4), 203-212.
Kelly, A. (2004). Design research in education: Yes, but is it methodological? *The Journal of the Learning Sciences*, *13*(1), 115-128.
Kelly, A. E. (2003). Theme issue: The role of design in educational research. *Educational Researcher*, *32*(1), 3-4.
Kelly, A. E. (2006). Quality criteria for design research. *Educational Design Research*, 107-118.
Kelly, A. E., Lesh, R. A., & Baek, J. Y. (2008). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* Abingdon, Oxon: Routledge.
Lester, F. K. (2005). On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *ZDM*, *37*(6), 457-467.
Margolis, J., & Fisher, A. (2003). *Unlocking the clubhouse: Women in computing.* Cambridge, MA: The MIT Press.
McKenney, S., & Reeves, T. C. (2013). Systematic review of design-based research progress is a little knowledge a dangerous thing? *Educational Researcher*, *42*(2), 97-100.
Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement*, *42*(5), 34-37.
Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, *40*(7), 331-337.
Pfeffer, J. (1993). Barriers to the advance of organizational science: Paradigm development as a dependent variable. *The Academy of Management Review*, *18*(4), 599. doi:10.2307/258592
Plomp, T., & Nieveen, N. (2007). An introduction to educational design research. In *Proceedings of the Seminar Conducted at the East China Normal University, Shanghai (PR china), November 23-26, 2007.* Enschede: Netherlands Institute for Curriculum Development.
Schoenfeld, A. H. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher*, *28*(7), 4-14.
Simon, H. A. (1996). *The sciences of the artificial (3rd ed.).* Cambridge, MA: MIT Press.
Stokes, D. (1997). *Pasteur's quadrant: Basic science and technological innovation.* Washington, DC: Brookings Institution Press.
Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training.* Abingdon, Oxon: Routledge.
Thompson, C. (2007). Halo 3: How Microsoft labs invented a new science of play. *Wired Magazine*, *15*(09).
Van den Akker, J. (1999). Principles and methods of development research. *Design Methodology and Developmental Research in Education and Training*, 1-14.
Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research.* Abingdon, Oxon: Taylor & Francis.
Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 493-502).

Acknowledgments

We thank Bruce Sherin, Pryce Davis, the Delta Lab, and the anonymous reviewers for their thoughtful critiques of this paper. This work supported by the National Science Foundation Grant No. IIS-1320693 and No. IIS-1217225.

Modeling the Construction of Evidence through Prior Knowledge and Observations from the Real World

Lauren Barth-Cohen, Daniel Capps, Jonathan Shemwell, University of Maine, Center for Research in STEM Education, 5784 York Complex #1, Orono, ME 04469
Email: lauren.barthcohen@maine.edu, daniel.capps@maine.edu, jonathan.shemwell@maine.edu

Abstract: Evidence is key to many scientific practices including argumentation. For learners engaged in scientific practices, we aim for them to recognize scientific evidence from observations in the natural world. Here, we provide an early depiction of evidence construction, namely *how* evidence is constructed from one's prior knowledge and one's observations. We illuminate instances of teachers constructing evidence while engaged in a professional development workshop where they are tasked to reconstruct the geological history of a national park. We illustrate four cases, some of which involve the successful construction of evidence and some of which involve embedded challenges with constructing evidence, such as difficulties with background knowledge and individuals "seeing" different information in the same phenomena. This analysis illustrates the role of prior knowledge in scientific practices that rely on evidence construction in field-based complex environments.

Introduction

A central focus of the Next Generation Science Standards (NGSS) is engaging students in the learning of scientific practices, such as argumentation and modeling. Here we focus on evidence construction, which we argue is important for many of these scientific practices. Science education research discusses the importance of evidence in the scientific process and in students' engagement in these practices in the classroom. However, less frequently discussed in this literature is research on *how* learners develop evidence from their immediate experience of phenomena. Possibly this issue is overlooked because often learners that are engaging in scientific practices are provided with evidence a priori. Here we use theoretical machinery from coordination class theory (diSessa & Sherin, 1998) to show how learners constructed evidence from their prior knowledge and observations of the world. The learners were middle school earth science teachers who observed the geology of an area and constructed evidence for the relative ages of the rocks and other features based on their observations and prior knowledge.

Evidence and Evidence Construction in Scientific Practices

Evidence is at the core of knowing and doing in science. No idea or explanation can be accepted if it fails to square with the "facts of the matter." Ironically, the origins of evidence construction—where the "facts" come from—are complicated. Fundamentally, evidence consists of straightforward observations dependent upon the five senses (i.e., perception). However, what is perceived depends upon the knowledge and expectations of the observer (Chalmers, 1999). To put it another way, the "facts" of any "matter" are never directly apprehended; they are constructed in the mind of the observer. Science educators acknowledge this fact along with its implications for learners engaged in scientific inquiry. For instance, Duschl (2003) described how learners begin with observation and end with explanation as they move through a series of transformations on an evidence-explanation continuum. The learner begins with the transformation from *data to evidence*, then goes from *evidence to patterns and models*, and finally moves from *patterns and models to explanations*.

Science education standards (e.g., NGSS, 2013) emphasize that evidence is paramount to student understanding of science. The 2007 National Academy Press report describes that students who understand science should be able to generate and evaluate scientific evidence (Duschl, Schweingruber, & Shouse, 2007). In the Framework for K-12 Science Education, a commitment to evidence is foundational for developing further claims and for engaging in many scientific practices, such as explanation, argumentation, and modeling (NRC, 2012). When constructing scientific explanations, using appropriate evidence is key for those explanations (McNeill & Krajcik, 2007). In argumentation, evidence is one of the pillars of the widely used Toulmin Argumentation Pattern (e.g. Erduran, Simon & Osborne, 2004). In modeling, the science content and relevant data is centrally important because a model needs to be *of* something (Lehrer & Schauble, 2006).

Construction of evidence from interactions with the natural world has rarely been the focus of research on the use of evidence in science education. In many studies, evidence has been provided to students as the starting point for the instruction process. Studies of this type have shown that students have difficulties both using appropriate evidence (McNeill & Krajcik, 2007; Sandoval, 2003) and understanding what counts as evidence (Sadler, 2004). These challenges are also prevalent in adults (Kuhn, 1991). McNeill and colleagues (McNeill et al., 2006; McNeill & Krajcik, 2007) conducted a series of studies in which learners constructed scientific explanations using claims, evidence, and reasoning. The data students used as evidence in their

 © ISLS

explanations were derived from either chemistry investigations or observations or reading materials. Bell and Linn (2000) presented the results of a student debate about how far light travels. The debate was scaffolded through an argument building software called SenseMaker where students were presented with 13 pre-existing items of evidence, but they had the option of adding unique evidence. From another perspective, research using a technology supported curriculum called Explanation Constructor had students construct and evaluate scientific explanations for natural phenomena, mainly the natural selection of finches in the Galapagos Islands (Sandoval, 2003; Zembal-Saul et al., 2002). Built into the curriculum was a database of information such as field notes about bird behavior and environmental factors (e.g. numbers of plants and animals in a given season). The students could query the data, and the software automatically generated graphs to scaffold the data analysis being easily connected to explanations. Studies like these have led to a line of research that focuses on developing and implementing curricula where students are asked to reason with evidence and ground their explanations and arguments in evidence (e.g. Zembal-Saul, et al., 2002).

Studies in the broader learning sciences literature have grappled productively with how learners might interact with phenomena to construct evidence. Goodwin (1994) proposed that "nature is transformed into culture" (p. 607) through processes of coding, highlighting, and representation. His point was that these processes, involving interactions with people and tools, organized learners' perceptions of nature to develop categories of importance needed for "practices of seeing" (p.608). He illustrated the processes with archeology graduate students learning to distinguish and interpret different colors and patterns in dirt. In an example of highlighting, an archeologist drew a circle around an area of colored soil to indicate to others where an ancient wooden post had been. Coding occurred when students learned to compare the colors of dirt on a trowel to those of a standard color chart, where the color of dirt on the trowel never quite matched any color on the chart. In a review of the literature, Mogk and Goodwin (2012) described how geology graduate students in a field camp learned to "construct and 'read' the story of the Earth" (p. 145). The authors described how students, guided by their professor, scrutinized minute changes in the size, shape and distribution of rock to select and make crucial "first inscriptions" (p. 145). However, the authors did not go deeply into how evidence construction occurred except to make the general point that it was mediated by the professor's guidance, tools, and interactions with peers.

Evidence construction in this type of natural world setting is important because there is a wide-ranging space of potential evidence that a learner needs to sift through and that sifting process may support learning. Geologists have long recognized the importance of field experience in geology education. Thus, it is not surprising that researchers in geocognition have begun studying learning in the field (e.g. Petrocovic, Libarkin, & Baker, 2009).

Studies in the literature on expertise have shown how experts construct evidence. For experts, constructing evidence is nearly automatic because they are so familiar with the phenomena they encounter. That is, experts' evidence is largely pre-constructed, in memory. This point is illustrated by Eberbach and Crowley (2009). Reviewing the work of Alberdi, Sleeman and Korpi (2000), Eberbach and Crowley described how botanists compared features of plants. The botanists fluidly identified the relevant parts of plants and their features in order to identify similarities and differences between different organisms. Similarly, Knorr-Cetina and Amann (1990) presented an ethnographic study of how visual evidence was constructed by a community of molecular geneticists who were discussing a film. The translation of this visual film into data and then evidence was again a fluid process involving jointly produced talk by the scientists as they were interacting with each other and the object in question. However, needed are studies that illuminate how this process of constructing evidence unfolds with learners.

Our data was collected as part of a multiple-day geologic field experience in which learners made observations of the bedrock, and those observations had the potential to be constructed into evidence for complicated historical events. The current study builds off geocognition literature on learning in the field by describing learners' construction of evidence as they developed models of historical geology. We present data from a teacher professional development (PD) workshop where participants were engaged in model development and model revision in the field. Although a PD workshop is not inherently a naturalistic setting, this course primarily occurred outside in a complicated field geology environment where there was a large span of phenomena that the teachers could focus on while constructing evidence. Thus, our research setting was a step toward the type of real world settings that have been less frequently focused on in the broader science education literature. Our research goal was to begin conceptualizing learning in this complicated field geology environment, where there was a wide range of approaches and reasons for how evidence was constructed.

The contribution of this paper is to show how the evidence construction process unfolds from the learners' sensory perception of phenomenon in natural settings to their conception of the phenomena as evidence in relation to a claim. Both perception of phenomena and conception of evidence are inherently knowledge-dependent (Chalmers, 1999; Metz, 2000), therefore, our model focuses on learners' knowledge as they construct evidence in this complex environment.

Data Collection Context

Data were collected during a three-day (17 hours) PD workshop for experienced sixth grade earth science teachers. The workshop focused on evidence-based modeling activities to support the implementation of the *NGSS* (2013). The workshop focus was reconstructing the geological history of the Schoodic Peninsula in Acadia National Park. During the workshop the teachers made observations and drew surface maps and cross-sections of the bedrock at three locations, while learning about geological principles. Over the course of the workshop, the teachers developed a series of increasingly sophisticated drawings, or models, of the Schoodic Peninsula at three points in time based on the geological features they observed. The oldest feature they observed was the granite bedrock of the area (~400 million years). Next were a series of ~200 million old diabase dikes. Diabase, an igneous rock similar in composition to basalt, forms as magma flows into fissures in surrounding rock (granite in this case) and cools. Figure 1 shows several teachers sitting on a large dike. Figure 2 shows a series of small dikes that are a few centimeters wide. The third set of features that teachers observed were remnants from the most recent glaciation (~20,000 years ago) including glacial erratics, which are rocks that have been transported by glaciers, and markings in the bedrock, such as chatter marks, striations, and cracks. Figure 3 shows an example of the striations that the teachers observed. Teachers' models from the first day of the workshop focused on diabase dikes and glaciation at Schoodic point, which is on Schoodic Peninsula. The models from the second day were more complicated due to additional observations of chatter marks, striations, and cracks from a new location on the same peninsula, Fraser point. The models on the third day were even more sophisticated as teachers added information about tectonic plates from geologic maps and simulations. The workshop instructor was one of the authors and there were two geologists and three education researchers, including another author, supporting the instruction and data collection.

Data consisted of video and audio recordings from both fieldwork and classroom instruction. Additionally, the data corpus included all drawings and notes created by the teachers and researchers field notes. Initially we open coded the data (Charmaz, 1995). Open coding led to focused codes related to geological content (e.g. diabase dikes, glacial erratics, chatter marks) and aspects of argumentation (e.g. rebuttal, claim).



Figure 1. Several of the teachers sitting on a diabase dike.



Figure 2. An example of a diabase dike.



Figure 3. An example of glacial striations on diabase.

Model of How Evidence Was Constructed from Observations of the Real World

We use coordination class theory to show how learners constructed evidence in natural settings from the interaction of human sensory data with prior knowledge. The use of coordination class theory was motivated by recognition that during the PD workshop the teachers explored a set of geological features which many of them had seen before (all teachers lived in the area and had visited the park in the past, many had even participated in field trips led by park naturalists), but teachers also commented about seeing new things ("It looks the same, but I'm noticing more though.") since the last time they visited the location. These comments raised questions about "seeing." What did it mean to "see" new things even if you have literally "seen" the same thing before? How did what someone "saw" relate to the geological models they constructed?

Coordination class theory is an outgrowth of Knowledge in Pieces, which is an epistemological framework and analytical perspective on learners' prior knowledge (diSessa, 1993; diSessa & Sherin, 1998). One research direction focused on the nature of knowledge structures that underlie conceptual understanding. Coordination class is a model of a particular type of concept that is common in physics. Prototypical examples include force and acceleration (diSessa & Sherin, 1998). Recently, the coordination class model has been applied to other concepts, including probability (Wagner, 2006) and oscillatory motion (Parnafes, 2007). Coordination classes were originally built to capture conceptual learning, but here we apply a subset of the theoretical machinery to evidence construction. We present what diSessa (1991) referred to as a mini-model, which is a type of localized theory, aiming to get at the essence of a phenomenon and address "why questions." In this case, we present a model for how teachers constructed evidence, based on their in-the-moment, observations of bedrock geology, in order to support subsequent claims.

According to coordination class theory, there are phenomena in the world that have the potential to be

"seen"—they can literally be *read out* [observed] of a situation. To read things out of a situation one needs a strategy, or a way of perceiving the phenomenon that determines what is attended to. This is known as a *readout strategy*. The observation generated by the readout strategy is the readout. There are many possible readout strategies, and they can involve perception with the eyes or other senses. The *causal net* is the entire knowledge system, composed of both knowledge and strategies that will help determine how the information—that is momentarily read out—relates to the overall knowledge system. When one uses readout strategies to read out [interpret] information in the world there is an instantaneous interaction between the information that is read out and specific *knowledge pieces* within their causal net. This instantaneous interaction is referred to as an *inference* as it functions as the link between the observable information and the determination of things that are not as easily observable. One or many inferences together become *evidence*. This model can be illustrated with an example: suppose an individual is making observations of a diabase dike that is cross-cutting granite while constructing evidence for a claim about the granite being older than the diabase. The individual might read out specific information about the granite cracking in linear patterns and the darker diabase exhibiting a circular or swirling pattern within those cracks (see Figures 1 and 2). Instantaneously, the information that is read out might be combined with prior knowledge into a subsequent inference about the molten diabase having flowed into the existing cracks in the solid granite. This inference (possibly combined with other inferences) then might function as evidence for a claim about the diabase being younger than the granite because the granite would have had to already exist when the diabase intruded. However, it is also possible that an individual might only read out the type of rock, granite or diabase, and then not go any farther to connect what they observe with any scientific claims. In that case they have made observations of the rocks, but they have not constructed evidence. Another possibility is that an individual might read out information and decide that it is unrelated. For instance, an individual might read out that lichen is growing on the rock, but their prior knowledge would guide them to instantaneously decide that lichen, which is biological, is unrelated to the claim about the rocks relative ages. In the moment their prior knowledge would attune those readouts such that some become inferences that then turn into evidence for subsequent claims and irrelevant readouts would be discarded.

We used coordination class theory to illustrate four cases of teachers constructing evidence from bedrock as they attempted to reconstruct geologic history of Schoodic Peninsula. In the first case, the teachers focused on cracks in granite where diabase crosscut the granite. This case illustrates how the teachers successfully constructed evidence for the relative ages of the rocks—granite being older and diabase being younger—based on the read out information as it is connected to their prior knowledge and subsequence inferences. Similarly, the second case illustrates a teacher successfully coordinating multiple sources of evidence into a series of claims. In the third case, teachers focused on marks in the bedrock that may or may not have been chatter marks. The teachers lacked sufficient background knowledge to make conclusive inferences about the chatter marks, and therefore were hindered in constructing evidence for glaciation. In the fourth case, several teachers discussed whether or not diabase dikes extended below the surface. Individual teachers had different answers to this question depending on what they read out in the situation. This case illustrates that individuals read out different information from the same phenomena, and these different readouts could lead to the same phenomena being construed as evidence for different claims. The third and fourth cases were similar as both identify challenges in constructing evidence.

Case I: Successful Evidence Construction about Diabase Dikes Crosscutting Granite

This is a case of unproblematic evidence construction as a teacher was able to draw on her prior knowledge with relative ease. This case came from the first day of the professional development workshop when two teachers, Barbara and Naomi spent 20 minutes discussing diabase dikes. Instructionally, we aimed for the teachers to first, see the diabase dikes as cutting through the granite (Figures 1 and 2) and second, use the geological principle of cross-cutting relationships to deduce that the diabase was younger than granite. Teachers were asked to construct surface and cross-sectional maps to support their thinking.

The teachers were reading out information about the granite cracking in relatively straight lines and connecting that information with their prior knowledge to make inferences about the diabase filling in the existing cracks in the granite. Barbara claimed that granite was older and diabase was younger ("I think what happened is that the granite was here first and that was the oldest and the black came up through the cracks from the mantle."). Meanwhile, the instructor asked the two teachers to consider the alternative, granite being younger and diabase being older. In response, Barbara constructed evidence from her observations for her claim that granite was older and diabase was younger. She read out information about the two rocks exhibiting different visual patterns: the granite appeared to crack or fractures in straight lines and the diabase had a curvy or flowing appearance (see Figure 2). Barbara's prior knowledge may have directed her to pay attention to these visual patterns, which were then instantaneously read out and became an inference that the diabase flowed into the granite. This inference, for her, was evidence in support of the claim that diabase was younger than granite.

Barbara: It just seems like the black [diabase] flowed around because, when you look at the stripes going up, like, this one here. To look at that, it looks like that kind of filled in a crack that was already in the granite. As opposed to, if it was going another way I don't see the black splitting like that, as well, because it looks like it doesn't split like that as much. When you look at the rock [diabase] itself, it doesn't not have those clean sort of fractures that you see in, the granite seems to have those.

Instructor: So, I hear you telling me that you are seeing the black stuff seems to cut into the granite.

Barbara: Yeah, looks very different // Instructor: So // I feel like it filled into a fracture that was there in the granite, the granite fractures nicely like that. Where as, I don't know if this [black rock] does as much, I mean this seems to flow around, in it's pattern too, it seems like it flowed around things and fills in cracks that were already in the granite, its kind of flowing in.

Barbara read out that the "granite fractures nicely" and generated an inference about the motion of the diabase when it was a melt. The diabase "fills in cracks that were already in the granite" and "it just seems like the black [diabase] flowed around." Diabase is a solid rock, but her inference was about its historical motion. Additionally, she read out information that refuted an alternative inference that the diabase did not crack in a way that was similar to granite. "When you look at the rock [diabase] itself, it does not have those clean sort of fractures that you see in, the granite seems to have those." Using the patterns she read out from the rock, combined with her prior knowledge, Barbara produced evidence to support her claim that granite was older. The evidence also refuted the alternative inference, that the diabase was older. These claims were also supported by another inference and relevant readout about the light colored rock as granite and the black rock as different from granite. In addition to the claim about relative ages of diabase and granite, there was also a proximate claim about diabase having flowed in the past.

In summary, Barbara read out information about (1) the granite "cracking," and (2) the diabase having a particular pattern with none of the clean cracks of granite. Barbara generated an inference that the diabase flowed around the granite. This inference was taken to be evidence for the claim that the granite was older than the diabase. This case illustrates successful construction of evidence based on the information that was read out and connected to their prior knowledge and supported inferences about the diabase flowing, which in turn functioned as evidence for the claim that granite was older than diabase.

Case II: Incorporation of Multiple Pieces of Evidence into a Model

This is a case of a teacher successfully incorporating multiple pieces of evidence into a model. At the end of the morning of the first day, the teachers were asked to construct models of what was happening geologically at three points in time based on prior observations. Barbara mentioned information that she read out from the rocks and several inferences functioned as evidence for her claim about granite being older than diabase.

Barbara: It appeared that the granite was kind of fractured, very straight lines. And that the, diabase, if it was it diabase, I have to find out about that. It appeared that the black rock [diabase], was probably the middle age rock, warm and molten, like it must have flowed through the cracks and crevices. And went around things. Looking at it, in places where it was swirling, from places where it was being cooled. So we thought it was middle. And then the glacial erratics, which we didn't see any in that area, but the one I had seen on the other side, was definitely. And, I had said it was the youngest rock, but then Naomi brought up, well said it was may not have been the youngest but newest arrival, to this area. So, it was somewhere else and then just brought here. So I think that was...

Gina: You don't know what it was made of?

Barbara: It looked like granite, but different. I don't know if it was Lucerne Granite, but I don't know. It was all grey and white as opposed to the pink granite that we were seeing.

Geologist: What about the very bottom there? What did you all…?

Barbara: Glacial erratic were sitting on top of both rock layers. It was just seeing like it was dropped there. You know, it didn't look like anything else, and it was this rounded rock that didn't fit in with sharp edges of granite that we were seeing around it.

Barbara discussed her evidence for granite being older than diabase. She read out information about the straight-line fractures in the granite, which supported her inference that the molten black diabase flowed through the existing cracks and crevices in the granite while it cooled. This inference then became evidence for the

diabase being middle aged (200 million years old) and the granite being older.

Then, Barbara switched to discussing glacial erratics. She mentioned having seen a glacial erratic, which suggests that she had read out the features of a particular rock and connected them with relevant prior knowledge in order to identify that rock as a glacial erratic. The glacial erratic was taken to be evidence of the youngest rocks; we do not know why, likely, she had prior knowledge of the recent glaciation period, which was discussed elsewhere in the workshop. Gina asked Barbara about the composition of the glacial erratic. Barbara mentioned that it looked like granite, which implies that she read out information from the rock and connected that to her prior knowledge to determine that the specific rock was granite, but she was unsure about whether or not it was a particular kind of granite, Lucerne Granite. This provisional identification was made from her having read out information about the color of the rock. It looked grey and white, instead of pink, which was the typical color of the surrounding rock. At the end, one of the geologists asked her about glacial erratics. She then generated an inference in which glacial erratics were sitting there, possibly dropped there. She may have read out information about how erratics looked different from surrounding rock; the surrounding granite had sharp edges while the erratic was rounded. Accordingly, Barbara constructed a series of claims, granite as the oldest rock, diabase as the middle aged rock, and glacial erratics as the newest arrival. In generating these claims she constructed evidence for why granite was older than diabase, it appeared that the molten diabase flowed into existing cracks in the granite, and why the glacial erratics, which were composed of ancient granite, were newer arrivals to the local area than the granite bedrock.

Case III: Problematic Evidence Construction due to Uncertainty about Chatter Marks

This is a case of problematic evidence construction as the teachers lacked sufficient knowledge to determine the nature of the markings they were observing. This case came from the second day of the workshop when Barbara and Naomi discussed marks in the bedrock that may or may not have been chatter marks. They made observations of glacial markings at Fraser Point in order to develop a second round of models for the three previously mentioned time periods. Based on their new observations, Barbara and Naomi were unsure of whether the specific marks they were "seeing" were chatter marks or not. Chatter marks are evidence of glaciation and have a characteristic quarter moon shape that is aligned with the direction of glacial travel. The two teachers discussed several alternatives, perhaps these marks were cracks, striations or scratches, and they discussed how they might know if these marks were chatter marks or not. This conversation occurred over six minutes, one small excerpt is presented below. Throughout the conversation the teachers wrestled with whether the different markings they were observing on the rock were chatter marks or not. The issue of chatter mark identification was not resolved. Therefore, this is a case of trying to construct evidence, but there was little success due to uncertainty in determining chatter marks.

Naomi and Barbara read out several markings in the bedrock and the information read out as combined with their prior knowledge influenced them when trying to decide what kind of markings they were. Naomi first read out information about a marking when she asked Barbara if the markings were striations ("I haven't seen actual striations in a long time, could this be?"). Barbara responded with a possible inference about chatter marks ("Looks like chatter marks, it could be thought, going in the right direction?") and included information about why these markings might be chatter marks: they are going in the right direction. Barbara drew on prior knowledge about direction being important and she read out both the marks and their direction. Naomi agreed with the directionality, but had a concern about what was causing the chatter mark ("Could be something else dragging it."). Then Barbara read out new information about the depth of the marking ("It could have been, doesn't look like it was that deep, you know.") At this point they had read out some markings that might be chatter marks, but had not made a solid inference about if these were or were not chatter marks. Then, the conversation switched to a different mark in the rock and Barbara again asked about the identification ("What did they say, about these here? These cracks? Did they say that they are striations?"). In response Naomi said, "I've seen striations before. I've had them pointed out to me by a geologist, and I didn't think. That looks more like a crack to me." Here she was explicit about her relevant knowledge, having had striations pointed out to her in the past, she was uncertain about whether these were striations. Barbara then agreed, and Naomi continued to explain why this one was not a chatter mark ("Cause this should be more like a scratch.") Again, the conversation switched to a different mark in the rock, and this time Barbara was more convinced of it being a chatter mark ("There is some down there. It seemed like it was better, looks like chatter marks.") Naomi and Barbara again read out information about directionality and decided that these were more likely to be chatter marks ("they chatter marks were going in that direction. Kind of like that. I bet that's what they are"). This conversation stretched on while they continued to observe and try to identify additional markings.

Both teachers accessed some prior knowledge about identifying chatter marks as connected to glacial directionality. They read out specific markings in the rock including details about depth and shape, which implied direction. These readouts were connected with relevant prior knowledge while they tried to generate an inference about whether or not specific markings were or were not chatter marks, and they also discussed alternatives. In a short time period they looked at several markings with varying degrees of confidence. The

inferences were tentative because the teachers were unable to decide if the markings were or were not chatter marks and they were unable to make a clear argument for their uncertainty. This tentativeness might have been a manifestation of a lack sufficient prior knowledge. The lack of a solid inference that could have lead to evidence is important because chatter marks are generally evidence of glaciation, and thus they did not construct evidence for glaciation in these moments. Another factor is that the specific markings at this location were not ideal cases of chatter marks, and we would expect this to cause ambiguity for learners. This case illustrates that construction of evidence can become problematic when one either does not have sufficient prior knowledge or a certain kind of prior knowledge or when the readouts are ambiguous.

Case IV: Different Readouts of Whether or not Dikes Extend Below the Surface

Similar to Case III, this is a case of a challenge with constructing evidence, namely conflicting readouts leading to different claims. This case took place on Day 1 of the workshop when the teachers were in the field making observations of the bedrock, and they constructed three models, one for each time period. The instructor aimed for the teachers to see the dikes as cutting through the granite and running for a significant distance horizontally. In this instance, a small group of individuals discussed the question of whether the dikes extended below the surface or not. One possibility was that dikes seen on the surface extended below the surface in vertical columns. Another possibility was that dikes were inclusions, or chunks of rock, stuck in the middle of the granite and did not extend below the surface. Settling this question was important because two different conclusions could be reached based on one's interpretation of the evidence. If the dikes were viewed as cross-cutting the granite this would be used as evidence of dikes being younger than granite. On the other hand, if the dikes were viewed as inclusions in the granite, than the diabase would be older than the granite.

The teachers were asked to draw cross-sections to show the relationship between the rocks below the Earth's surface. Initially, teachers had trouble drawing what was below the surface. The instructor suggested the teachers should find a place where the rock was cut so they could see a cross-sectional view beneath their feet. The instructor directed the group to look below and asked if the dikes extended below the surface. ("You might have to look around a bit. But you might find a place where it was cut and you can look below. But, would you assume, like, coming over and looking at this. Would you assume that this is going to be really, really deep?") In response he heard both "yes" and "no," which prompted the next conversation about whether the dikes extended below the surface or not. Gina was not reading out information about the dikes going underneath ("You don't see if from down there, there isn't anything that comes out from underneath.") This supported the inference that the dikes did not extend below the surface. Lisa and Kelly both read out information about the dikes that supported a different inference: dikes extending below the surface. "Lisa: No, but it could go down there. It is going down there instead, so it's vertical. That doesn't mean that it doesn't stand up and (muffled argument). Kelly: Could be a big column." Then, Carol, similar to Gina, also read out the dikes as a chunk ("I think of it as a chunk of") and inferred that it did not extend below the surface.

This case is interesting because the teachers were likely all looking at the same rocks, but not reading out the same information. From different readouts we would expect them to draw different inferences, and construct different evidence. Lisa and Kelly generated an inference that the dikes continued below the surface, which could be taken as evidence for a claim about dikes being younger as they cut through the granite. Gina and Carol generated an inference that the dikes did not continue below the surface, which led to different inferences, evidence, and claims than Lisa and Kelly. Another possibility that cannot be ruled out is that they were drawing on different prior knowledge, however, given that the teachers appear to "see" different things from the same apparent rock, there is an indication of them having different readouts and not only different knowledge. Reading out the pertinent information is important in evidence construction, and this case suggests that one challenge in constructing evidence is what information one reads out of a situation.

Discussion and Conclusion

In this paper we used coordination class theory to show how learners constructed evidence in the field as their prior knowledge interacted with their observations. Cases one and two illustrated successful construction of evidence while cases three and four each illustrated a challenge with constructing evidence, including difficulties with background knowledge, and individuals having conflicting readouts. Coordination class theory was useful in describing these examples because it provided a way to conceptualize successful evidence construction and what can be problematic, both of which may impact how individuals learn. As we pointed out in our review of the literature, evidence construction is currently under-conceptualized. Thus, our work provides an early depiction of what evidence construction might look like in a field-based setting. Current research in geoscience education recognizes the importance of field experience. Studies on expertise have shown how experts construct evidence, but our work goes a step beyond this by beginning to elucidate how evidence is constructed in a field setting. We also see our work as generally applicable to the science education community. Evidence construction is a major part of many of the scientific practices highlighted in the recently released *NGSS*. Thus, understanding both the process and the difficulties learners might encounter in evidence

construction should inform the development of heuristics that can support learners in successfully constructing evidence as they engage in the learning of science.

References

Alberdi, E., Sleeman, D. H., & Korpi, M. (2000). Accommodating surprise in taxonomic tasks: The role of expertise. *Cognitive Science, 24*(1), 53-91.

Bell, P. & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797-817.

Chalmers, A. F. (1999). *What is this thing called science?*. Univ. of Queensland Press.

Charmaz, K. (1995). Grounded theory. In J.A. Smith, R. Harre, & L. Van Langenhove (Eds.), Rethinking methods in psychology (pp. 27–49). London: Sage

diSessa, A. A. (1991). If we want to get ahead, we should get some theories. In *Proceedings of the 13th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 220-239).

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*(2-3), 105-225.

diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*(10), 1155-1191.

Duschl, R. A. (2003). Assessment of inquiry. *Everyday assessment in the science classroom*, 41-59.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and Teaching Science in Grades K-8.* National Academies Press.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research, 79*(1), 39-68.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*(6), 915-933.

Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606-633.

Knorr-Cetina, K., & Amann, K. (1990). Image dissection in natural scientific inquiry. *Science, Technology & Human Values, 15*(3), 259-283.

Kuhn (1991). *The Skills of Argument*. Cambridge University Press, Cambridge.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. *Cambridge Handbook of the Learning Sciences,* 371-388.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences, 15*(2), 153-191.

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. Thinking with Data, 233-265.

Metz, K. E. (2000). Young children's inquiry in biology: Building the knowledge bases to empower independent inquiry. In J. Minstrell & E. H. van Zee (eds.), *Inquiry into Inquiry Learning and Teaching in Science* (pp. 371-404). Washington DC: American Association for the Advancement of Science.

Mogk, D., & Goodwin, C. (2012). Learning in the field: Synthesis of research on thinking and learning in the geosciences. *Geological Society of America Special Papers*, 486, 131-163.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: The National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press.

Parnafes, O. (2007). What does "fast" mean? understanding the physical world through computational representations. *The Journal of the Learning Sciences, 16*(3), 415-450.

Petcovic, H. L., Libarkin, J. C., & Baker, K. M. (2009). An empirical methodology for investigating geocognition in the field. *Journal of Geoscience Education, 57*(4), 316-328.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513-536.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *The Journal of the Learning Sciences, 12*(1), 5–51.

Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction, 24*(1), 1-71.

Zembal-Saul, C., Munford, D., Crawford, B., Friedrichsen, P., & Land, S. (2002). Scaffolding preservice science teachers' evidence-based arguments during an investigation of natural selection. *Research in Science Education, 32*(4), 437-463.

Acknowledgements

This work has been supported by NSF Grant no. DUE-0962805.

Automatic Coding of Questioning Patterns in Knowledge Building Discourse

Jin Mu, Jan van Aalst, Carol Chan, Ella Fu
The University of Hong Kong, Pokfulam Road, Hong Kong
jinmu@hku.hk, vanaalst@hku.hk, ckkchan@hku.hk, ellafu@ied.edu.hk

Abstract: We propose a novel method for identifying questioning patterns, which are assumed to be one of the essential factors indicating the quality of knowledge-building discourse. The underlying principle of the proposed method is to extract syntactic and sematic information before segmenting the raw data and annotating them according to a multi-layer framework called ACODEA. As a bottom layer of the framework, the "pre-coding" phase makes it possible to translate the raw data into machine-readable and context-independent language, and to make Natural Language Processing tools aware of users' preferences and underpinning mechanisms of identifying the desired pattern. Explorative but promising evidence is reported toward a more comprehensive perspective by combining qualitative and quantitative methods to analyze the discourse data. Given those findings, we argue in favor of mixed methods of content analysis and they further generated directions for future methodological development and empirical applications.

Introduction

In computer-supported collaborative learning (CSCL) environments learners often communicate with each other via text-based, digital discussion boards (Rosé et al., 2008), and this has been argued to reflect socio-cognitive processes of knowledge construction (Vygotsky, 1986). During collaborative learning activities, individual learners interact with each other in a dynamic way, making it very difficult to measure and assess learning effects independently. This may be one reason why the focus of collaborative learning research has shifted from studying learning outcomes and products to studying learning processes (Dillenbourg, Baker, O'Malley, & Blaye, 1995). With an interest in the collaborative learning process, the focus has recently shifted again – this time from analyzing individual learning processes toward identifying collaborative patterns that positively influence learning. This shift is fundamentally grounded in our understanding of collaborative learning from socio-constructivist perspectives.

Although uncovering findings related to how collaborative knowledge creation is accomplished is useful, analyzing a huge body of discourse data manually is an arduous task that consumes much time and slows down the research progress substantially. Over the past decade, there has been a substantial effort to develop innovative technologies that enable automatic content analysis in the domain of CSCL. These techniques enhance the ability of traditional approached to extract patterns that are assumed to be essential in the cognitive and social processes of learning. Against this background, by using a Natural Language Processing (NLP) tool called TagHelper (Dönmez, Rosé, Stegmann, Weinberger, & Fischer, 2005) and its successor SIDE (Mayfield & Rosé, 2010), a multi-layer framework called ACODEA (Automatic Classification of Online Discussions with Extracted Attributes, Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012) has been shown to be optimized for fully automatic segmentation and context-independent classification of the desired patterns—e.g., the quality of argumentation in a text-based CSCL discourse data. By extracting syntactic and semantic features during a pre-processing phase before content analysis, the framework allows a bottom-up specification of the in-depth information contained within the discourse corpus and it is therefore more precise and reliable than traditional approaches. The goal of the present study is to extend the previous work on automatic content analysis by applying the ACODEA framework to data from Knowledge Forum. The on-going efforts herein are assumed to extend the capabilities of the classification models with the outlined steps to be quickly customized for different contexts and alterative coding dimensions of interest in the field of CSCL.

Knowledge Building and Automatic Content Analysis

Knowledge building refers to the development of innovative and sustained knowledge within a community (Scardamalia & Bereiter, 1994, 2006). The major objective of this pedagogical approach is to initiate students into a knowledge-creating civilization by encouraging them to engage in sustained idea improvement and advance the knowledge collectively as a community (Scardamalia & Bereiter, 2006; Zhang, Scardamalia, Reeve, & Messina, 2009). Consequently, it turn out to be essential to conduct content analysis which is capable of revealing what is developed through the continuous process of idea improvement and knowledge advancement at both individual and collective levels.

In the last decade, while various assessment approaches have been developed so intensively that some of the tools have been even integrated with Knowledge Forum - a technology-mediated learning environment to

 © ISLS

foster knowledge building - it has proven challenging to grasp the overall picture of the community-based learning process. In fact, the majority of current automatic approaches are still in early stages of development; previous research has mostly focused on detection of simple patterns rather than in-depth content analysis of discourse data. For instance, the Analytic Toolkit (Burtis, 2002) provides summary statistics on student participation and interaction in Knowledge Forum databases, by counting the instances of note creation, note reading, and note linking. Similarly, applet tools (Zhang, Hong, Scardamalia, Teo, & Morley, 2011) for social network analysis (SNA) have been used to explore the social structure of collaborative discourse by offering quantitative indices, such as network centrality in networks based on reading behaviors (i.e., who has read whose notes). However, little attention has been given to the quality of knowledge advancement and reflection on the depth of cognitive and social processes taking place during the collaborative learning. Recently, van Aalst et al. (2012) took one step forward to explicitly analyse the quality of knowledge-building discourse by developing a tool for formative assessment – the Knowledge Connections Analyzer (KCA). The KCA was designed to create a model for the collaborative and epistemic patterns of collaborative knowledge construction by retrieving evidence on four key questions: 1) Are we a community that collaborates? 2) Are we putting our knowledge together? 3) How do our ideas develop over time? And finally 4) What is happening to my own ideas? Using this model, van Aalst et. al. (2012) began to illustrate the collective (Q3) and individual (Q4) aspects of idea improvement by extracting key words which were used most frequently to trace the awareness and use of new concepts appeared in the database.

Adapted from previous efforts within knowledge-building communities to conduct qualitative content analysis either manually or automatically (Carol Chan & Lam, 2010; van Aalst, 2009; Zhang et al., 2011), in the current study we intended to go beyond the existing approaches to further identify critical features of knowledge-building discourse by using advanced NLP technologies. Briefly, the NLP tool SIDE can automatically extract features like line length, unigrams, bigrams and part-of-speech bigrams from the annotated data to build models (Mayfield & Rosé, 2011). The process is similar to linear regression that expresses the classification categories as a linear combination of the attributes (extracted syntactic or semantic features) with predetermined weights (coefficients). We assume that the appropriate value of the predicted weights is dependent on the importance of the extracted features to reflect on the underlying epistemic and collective aspects of knowledge-building discourse instead of the simple accounting frequency.

Questioning in Knowledge Building Discourse

Questioning is a core function and a key feature of both learning and teaching, and good questions can stimulate students to think at higher cognitive levels (Dillon, 1988). Furthermore, the questioning behavior in learning has consistently elicited elaborated explanations, inferences, justifications, speculations, and other essential signs of complex knowledge construction (King, 1994). While asking and answering questions are among the most common human activities, it is remarkable how little is known systematically about questioning , especially about the methods for measuring and analyzing the desired questioning patterns in CSCL.

It has been reported that over 75% of the questions posed in both elementary and secondary classrooms are “recalling” questions (Dillon, 1988). Approximately 3.5% of the questions are asked to check for understanding of procedures, routines, and only slightly more than 1% questions are at a higher cognitive level, such as evaluation and synthesis questions (Craig & Cairo III, 2005). In addition, learners are rarely observed to ask self-generated questions of the teacher or other peer pupils. Hence, the majority of studies in researching the effectiveness of questioning focus on teacher-generated questions and examine the relationship between such questioning behaviour and student achievement (Craig & Cairo III, 2005). However, when learners engage in knowledge-building discourse, in which learners play more central roles (Scardamalia & Bereiter, 1994), it would be useful to know whether students can generate higher-order questions, which lead students to think, analyse and synthesize the discussion topic at higher cognitive levels.

Craig and Cairo III (2005) identified six types of questions: Recall (facts from memory); Check for understanding of procedures and routines; Use (using knowledge to comprehend, apply, or analyse); Teacher repeats the question two or more times; Create (synthesizing to arrive at a conclusion) and Teacher asks multiple questions. According to King (1994), we need to differentiate "memory" questions which refer to those requiring learners to simply remember and repeat what they had heard and memorized from the lesson and "thinking" questions. The latter ones require learners to not only remember information from the lesson but also think about that information. Thinking questions were further classified into comprehension questions and connection questions. King (1994)) stated that comprehension questions “check how well you understand the lesson” and “ask you for a definition in your own words or ask you to tell about something you learned about-but in your own words, not the teacher's words” (p.346). Connection questions are thought provoking because they require students to go beyond what was explicitly stated in the lesson by linking two or more ideas together in some way. As a result, during a discussion the learners tended to make those connections between and among ideas, which may reflect the mental representations they constructed the links in mind. Such highly elaborated and richly integrated questions could account for the improved comprehension of the instructional material.

Learners have been regarded as being capable to ask and recognize two types of questions, namely text-based question promoted by text and higher-order knowledge-based questions stimulated by event (Carol Chan, Burtis, Scardamalia, & Bereiter, 1992; Scardamalia & Bereiter, 1992). In line with the previous research on knowledge building, questioning patterns have also been classified using two categories determined by the cognitive goals: "fact-seeking" and "explanation-seeking" questions. Explanation-seeking questions are embedded in the process of inquiry by asking "why" and "how", whereas fact-seeking questions are looking for "fragmented pieces of knowledge" (Hakkarainen, 2003, p. 1075). In another study (Lee, Chan, & van Aalst, 2006), further differentiated questions based on the nature of the information sought: 1) definitions and simple clarifications; 2) factual, topical and general information; 3) specific gaps in terms of open-ended responses and different viewpoints; and 4) explanation-based questions that focus on problems instead of topics and identify sources of inconsistencies; generates conjectures and possible explanations.

Three functions of question can, therefore, be identified in the present study. Simple statements of information or facts gleaned directly from the lesson, prior knowledge, or experience are coded as fact-seeking questions (e.g., "What is meant by zone of proximal development?"). Thinking questions (e.g., "What is the role of assessment in a learning community?") ask for deeper understanding of by translating into a student's own words and they are often elaborated upon by connecting with other conceptual ideas. Using questions that integrate aspects of the contextual information outside the learning environment assume to go beyond other question functions in some manner. An Example is "How can we make use of Knowledge Forum and really help the students to construct knowledge through reading?"). This kind of questions assume to effectively prompt students to connect learning content with their prior knowledge and personal experience with the purpose of resolving the authentic problems raised from real contexts. As mentioned above, a major concern of CSCL research focuses on in-depth analysis of collaborative learning processes. In the following, we will present an advanced approach to automate the content analysis of questioning behavior in knowledge-building discourse. The main question addressed in this study is: How does the ACODEA framework perform in automatically analyzing knowledge-building discourse data? We divided this question into three sub-questions: (a) to examine the reliability of capturing key patterns of questioning behaviours in an automatic way, (b) to explore the function and degree of questioning patterns in a knowledge-building community by applying the developed approach of automatic content analysis, and (c) to determine the effects of this automatic content analysis by comparing with other automatic approaches integrated in Knowledge Forum.

Research Questions

RQ1: *Can the automatic content analysis be implemented reliably to extract the key patterns of questioning behavior in knowledge-building discourse?* We expected to achieve an acceptable level of agreement between automatically generated codes by SIDE and human codes when we automate the text classification on the multi-layer ACODEA framework by extracting the desired attributes of questioning behavior in a systematic way.

RQ2: *Which function and degree of questioning behaviour would be more often exhibited in the knowledge-building discourse?* To answer this explorative question we needed to describe the frequency, type and quality of the questioning behaviors coded through the automatic approach investigated in RQ1.

RQ3: *To what extent the results of the automatic content analysis are related to those results reported by other automatic measurement approaches, such as the Analytic Toolkit (ATK)?* We hypothesized that notes embedded with higher-order questions are expected to be more widely read and built-on during the discussion. By examining the features of such notes, we hoped to gain some insight into why some of notes have more impact than others to be read or receive more build-on notes. In other words, we mainly concern how much variance in the number of reading and building-on can be explained by the extracted patterns of questions.

Methods

Participants and Learning Task

The participants consisted of more than 40 teachers, researchers, and graduate students who were part of the BCHK Network in 2002 (CKK Chan & Van Aalst, 2003). The Knowledge Forum database for this networks hosted a course on knowledge building, but also contained online discourse of teachers who were attempting to implement other higher-order thinking strategies in classrooms, in line with a recent curriculum reform in Hong Kong that emphasized "learning how to learn" (CDC, 2001). Participants were required to contribute to online discussion on Knowledge Forum, which mainly focused on a set of independent but closely connected topics to acquire deeper understanding of knowledge-building and related theories, classroom implementation, the role of teacher, and instructional designs.

Data Source and Coding Processes

Altogether, there were 1742 notes and 65,535 words in the corpus collected from 5 Knowledge Forum views. Two human coders analyzed almost all of the raw data. About half of the human-coded data were used as the training materials on which a few automatic models can be built by SIDE (Mayfield & Rosé, 2010). The resulting model could then be easily applied to classify un-annotated data, and then the assigned codes could be further reviewed on the annotation interface that facilitates the process of humans correcting errors made by the automatic coding. The remaining manually coded dataset were further used for testing the training models. SIDE employs a consistent evaluation methodology referred to as 10-fold cross-validation, where the data for training the models can be randomly distributed into 10 piles. Nine piles are combined to train a model. One pile is used to test the model. This is done 10 times so that each segment is used as a test set once. And then the performance values are averaged to obtain to final performance value (Rosé et al., 2008).

By following the Automatic Classification of Online Discussions with Extracted Attributes (ACODEA, Mu et al., 2012), the coding process implemented in the present study consists of three layers. The general idea underlying the multi-layer framework of automatic content analysis is to extract features at the lower layer that assume to contribute to the text classification at the upper layer. For instance, a unit of analysis can be identified as fact seeking question (at the upper layer) by combining both contextual facts and question words (at the lower layer). (i) Regarding the semantic attributes extracted at the lower layer, each single word in the text was separated into one of the following categories: (a) Core Concept, keywords from knowledge-building theory and principles; (b) Peripheral Concept, keywords form relevant theories and learning sciences; and (c) Contextual Information from the learning environment and local settings. In addition, there were other attributes being of importance in reflecting the (d) Question Words and (e) Thinking Verbs as the key indicators of higher-order questions that were distinct from other (f) General Verbs. Examples of the extracted attributes are illustrated in the Table 1. (ii) The unit of analysis was defined as a sentence or part of a compound sentence that can be regarded as "syntactically meaningful in structure" (Strijbos, Martens, Prins, & Jochems, 2006). For instance, according to these rules of segmentation, punctuation and the special symbol like question mark are boundaries that can be used to segment compound sentences if the parts before and after the boundary are 'syntactically meaningful'. The segments can be further identified either as statement or question. (iii) The last coding layer was designed to capture the patterns of questioning behavior in knowledge-building discourse. In the present study we are mainly concerned with the categories (summarized in Table 2 below) by using a machine-readable coding system that captures the function of the questions ranging from low to high in complexity, and roughly corresponding to the two degrees for each function of questions.

Table 1 Extracted attributes and examples

Extracted attributes	Example
Core Concept	Knowledge Building, Principle, Collective, Expert, KF
Peripheral Concept	Ability, Gifted, Pedagogy, Notes, Views
Contextual Information	Hong Kong, Primary School, Mathematics, Physics
Question Words	Why, How, Where, When, Who, What, What if
Thinking Verb	Think, Wonder, Reflect, Test, Hypothesize
General Verb	Do, Have, Be, Can

Table 2: Coding schema to analyze the questioning patterns within knowledge-building discourse

Function of Questions		Degree of Question	
Fact-seeking Questions	Requiring learners to simply recall and repeat explicit and factual information	Low-level Yes or No Trivial Question	by asking for yes-or-no responses in terms of trial facts with brief wording
		High-level Open-ended Question	by asking for open-ended responses
Thinking Questions	Requiring learners to think about conceptual ideas for deeper understanding	Low-level Elaboration Question	by clarifying, elaborating and explain a conceptual idea
		High-level Connection Question	by linking two or more conceptual ideas together
Using Questions	Requiring learners to apply conceptual ideas to certain real context with the purpose of practice	Low-level Utilization Question	by applying peripheral concept to context
		High-level Application Question	by applying core concept ideas to context

Measuring Reliabilities of Automatic Content Analysis and other Variables

The reliability of the coding was measured using Cohen's Kappa value and percent agreement. Both of the indexes have been regarded as widely used standards for measuring coding reliability. Percent agreement is the most simple and most popular reliability coefficient (De Wever, Schellens, Valcke, & Van Keer, 2006).

Statistically, the inter-rater agreement is determined by dividing the number of codes that are agreed upon by the total number (agree and disagree all inclusive) of codes. Supplemental criterion for success is reaching a level of inter-rater reliability with a gold standard as measured by Cohen's Kappa that is .7 or higher (Strijbos et al., 2006). Here it is worthwhile to further clarify that the present study was undertaken to evaluate different types of Kappa including (1) inter-rater agreement between human coders Kappa (Human-Human) to evidence the initial reliability of training examples; (2) inter-rater agreement generated by the 10-fold cross-validation to certify the internal reliability of the SIDE training models. The 10 results from comparing the coding between SIDE and manually coded training materials then can be averaged to produce a single estimation Kappa (SIDE-Training); and finally (3) the conclusive Kappa (SIDE-Testing) between SIDE and human coders calculated with the additional testing materials.

With respect to other variables measured in the present study, both of the categorical variables *Function of Questions* (Fact-seeking, Thinking vs. Using) and *Degree of Questions* (Low vs. High) were coded by applying the approach of automatic content analysis developed for the present study. Another analytic tool Analytic Toolkit (ATK) that is integrated within Knowledge Forum provided information for reflecting on the *Number of Reading and Building-on* which refer to how many times the notes were read or replied by other members within the Knowledge Building community during the online discussion.

Results

Two different analyses were conducted in the present study. First, reliability of the various coding categories in the multi-layer framework was calculated and table displayed. Second, linear regression analyses were conducted to assess the degrees of association between automatically coded questioning behaviours and the number of reading and building on the notes as assessed by ATK.

RQ1: Can the automatic content analysis be implemented reliably to extract the key patterns of questioning behavior in knowledge-building discourse? Two coders created the training material for SIDE. The overall value of kappa on segmenting and identifying questions was statistically highly significant; Cohen´s Kappa (Human-Human) was 1.00 with 100 percent agreement that indicated a good degree of inter-rater reliability beyond chance. Additionally the human coders achieved a high value of Cohen's Kappa (Human-Human) = .89 (Percent Agreement = 91.7%) for the final coding layer. These results indicate acceptable human baseline performances for SIDE to be trained to analyze the un-annotated data regarding the extracted attributes, segmentation and coding layers.

SIDE achieved an internal Cohen's Kappa (SIDE-Training) = .73 (Percent Agreement = 96.7%) on the layer of segmentation. The reliability comparing SIDE with a human coder (based on raw text) was sufficiently high (Cohen's Kappa (SIDE-Testing) = .71; Percent Agreement = 89.0%). As shown in Table 3, sufficient inter-rater agreement values were also achieved for the second layer with Cohen's Kappa (SIDE-Training) = .94 (Percent Agreement = 99.1%) and Cohen's Kappa (SIDE-Testing) = .96 (Percent Agreement = 99.4%). Internal Cohen's Kappa (SIDE-Training) = .73 (Percent Agreement = 82.9%) was achieved by SIDE when it attempted to automatically code the questions with respect to the function and the degree. A human coder and SIDE achieved an agreement of Cohen's Kappa (SIDE-Testing) = .77 (Percent Agreement = 85.5%).

Table 3: Reliability of the multiple layers of automatic content analysis

Multiple Layers of Automatic Content Analysis		Cohen's Kappa	Percent Agreement
Layer i	Segmenting		
	Training (SIDE)	0.73	95.7%
	Testing (Human vs. SIDE)	0.71	89.0%
Layer ii	Identifying Questions		
	Training (SIDE)	0.94	99.1%
	Testing (Human vs. SIDE)	0.96	99.0%
Layer iii	Coding Questions		
	Training (SIDE)	0.73	82.9%
	Testing (Human vs. SIDE)	0.77	85.5%

RQ2: Which function and degree of questioning behavior would be more often exhibited in the knowledge-building discourse? Upon initial impression, there were 3465 single segments in total, and 263 of them were identified as questions. The results indicate that the community members did generate a number of questions spontaneously. Among them, slightly less than half (44.7%) of questions generated in the Knowledge Building discourse were thinking-oriented, only 28.2% of questions were seeking for factual information, and a rather low percentage of 15.3% linked to the using questions. The frequency percentage of various degrees of questions did not appear to be significantly different cross three functions. But participant appeared to be able to generate higher-order questions. For instance, in the questions asking for factual knowledge, roughly two third

of them were open-ended. The most frequently asked question was connection question at 25.2%, followed by elaboration questions at 19.5%. Perhaps not surprisingly, the participants tended to be more often to apply the knowledge-building theory at the lower level, given that 15.3% of the generated questions were classified as the utilization questions, followed by the higher degree of application questions (11.8%). The main patterns of the questioning behaviour in the KB discourse are summarised in Table 4.

Table 4: Frequency of various functions and degrees of questions

Questions Categories		Frequency	Percentage
Fact-seeking Questions		74	28.2%
Low Degree	Yes or No Question	*26*	*10.0%*
High Degree	Open-ended Question	*48*	*18.3%*
Thinking Questions		117	44.7%
Low Degree	Comprehension Question	*51*	*19.5%*
High Degree	Connection Question	*66*	*25.2%*
Using Questions		71	27.1%
Low Degree	Utilization Question	*40*	*15.3%*
High Degree	Application Question	*31*	*11.8%*
Total		262	100%

RQ3: To what extent the results of the automatic content analysis are related to those results reported by other automatic measurement approaches, such as the Analytic Toolkit (ATK)? A multiple regression analysis was performed between the dependent variables (separately, the frequency of Build-on and Reading Notes) and the independent variables (simultaneously, the Function of Questions in terms of two dummy coding variables Thinking and Using, the Degree of Questions Low vs. High, and the Authority of the Authors Researchers vs. Teachers). Analysis was performed using SPSS Linear Regress.

Table 5: The number of reading and building-on by other community members

Function of Q		Fact-seeking Questions		Thinking Questions		Using Questions		Total
Degree of Q		Low Degree	High Degree	Low Degree	High Degree	Low Degree	High Degree	
Reading	*Mean*	28.58	24.08	25.82	44.27	14.88	28.13	29.03
	SD	20.92	15.82	16.23	39.09	8.61	22.33	26.13
Building On	*Mean*	1.00	1.06	1.00	1.23	0.45	0.61	0.94
	SD	1.06	2.04	1.11	1.32	1.11	0.99	1.38
N		26	48	51	66	40	31	262

Regression analysis revealed that the model significantly predicted the number of reading and building-on by other community members. The model using the 4 predictors explained about 36.5% of the variance of the number of reading by others, F (4,257) = 36.89, $p < .001$. The predictor Thinking had a significant positive effect on the number of reading, $\beta = .14$, $p < .05$. Notes containing higher order thinking question were more often read than notes with fact-seeking questions, while another dummy coding factor Using was not a significant predictor of the number of reading during the online knowledge building discussion. $\beta = -.04$, $p > .05$. The Degree of Questions had a significant regression coefficient, $\beta = .16$, $p < .01$, indicating that notes with higher degree of questions were expected to be more read. Meanwhile, the Authority of the note author was of similar magnitude to predict the number of reading, $\beta = .52$, $p < .01$. Not surprisingly, researchers still provided focus to the discussion by posting more impactful notes read by other participants.

The R square value of the Build-on Model was lower, which was able to account for 19.7% of the variance in the model, F (4,257) = 15.78, $p < .001$. Different form the model of reading, the factor of Using had a significantly negative effect on the number of building-on, $\beta = -.15$, $p < .05$. Notes with the questions asked for utilization and application on authentic problems were expected to be surprisingly less desired regarding the number of building-on than notes with fact-seeking questions. Different from reading a note, building-on other's note is a more active and challenging task, especially when it is required for applying the core conceptual ideas to real context. Hence, teachers engaged in the knowledge building discourse shunned to respond to more difficult questions. Other variables did not contribute to the model.

Conclusions and Future Work

This study found promising evidence that questioning patterns in Knowledge Forum databases can be coded automatically using the ACODEA framework acceptable reliability. Moreover, although previous research only

evaluated the reliability for argumentation data (Mu et al., 2012), this study suggests that the method can also be applied to data created within a different theoretical framework—knowledge building. It suggests that the framework can be applied successfully for automatic content analysis on different construct of interest and crossing different domains. The particular strength of the method lies in the clear understanding of how the discourse properties of interest manifest themselves via a variety of linguistic terms (either syntactic or semantic), which can be further viewed as a natural extension of keywords targeted at the machine-readable and context-independent language to build text classification algorithms that are consequently more powerful than text classification directly based on raw data.

We also presented two methods for assessing a real-world data set of knowledge-building discourse. While qualitative content-based analysis appeared to be more effective to detect and analyze the desired discourse patterns than quantitative analysis of counting the reading and building-on behaviors in a superficial manner, when used in isolation the methods may not identify all of the aspects of the Knowledge Building discourse. For this reason, we combined both of the qualitative and quantitative methods to provide a full picture of what happened during the online discussion. Our classifiers can reliably identify multiple patterns of questioning behavior, which have been further shown to be able to explain and predict if notes can be more read or built-on as assessed by the analytic tool (ATK). In this way, integrating results from different evaluations into a global consideration brought new insight for us to analyze the discourse data comprehensively and deeply.

We now briefly discuss some avenues for future work. As one of the major contribution of the present study, automatic analysis not only intends to speed up research projects, it also brings insights into essentially changing the way how teachers and educators design learning environments and scaffold the desired collaborative learning. Specifically, automatic analysis of online discussion can provide instructors with the capability to monitor the real-time learning progress occurring in large classes, indicate what the specific and personalized need should be addressed and consequently enable the adaptive intervention, which is assumed to be more efficient in promoting productive collaboration and knowledge building, than the static, one-size-fits-all scaffolds (Gweon, Rosé, Carey, & Zaiss, 2006; Kumar, Rosé, Wang, Joshi, & Robinson, 2007; Stegmann, Mu, Gehlen-Baum, & Fischer, 2011). Practically, integrating the automatic assessment in Knowledge Forum can be of valuable assistance for teachers to get to know how well their students are learning with a much lower investment of efforts. Therefore teachers can scaffold individuals and groups of learners more effectively in formative assessment. Based on the current study, the implementation of a well-controlled, randomized experiment is needed to examine the efficacy of the automatic content analysis as an effective formative assessment technique.

In addition the newly developed approach seems to be promising to develop domain insensitive coding schemas to model similar behavioral patterns occurs knowledge-building discussion. In other words, it enables researchers to address the urgent need for the re-use of coding schemas in diverse contexts. While being a well-established tradition to reanalyze quantitative data in social sciences, conducting secondary analysis of qualitative resource collected by other researchers, e.g. text-based discourse is relatively scarce in the field of CSCL. Hence, the general goal of the preliminary investigation aims at developing a feasible model in a manner that allows content analysis focusing on context-independent perspectives. At the same time, we will try to promote the discussion among the researchers within knowledge-building communities to facilitate the secondary analysis cross various learning settings.

References

Burtis, J. (2002). *Analytic toolkit for Knowledge Forum*. Paper presented at the Institute for Knowledge Innovation and Technology, Toronto, Ontario, Canada.

CDC. (2001). Learning to learn - the way forward in curriculum. Hong Kong, SAR, China: Government Printer.

Chan, C., Burtis, P. J., Scardamalia, M., & Bereiter, C. (1992). Constructive activity in learning from text. *American Educational Research Journal, 29*(1), 97-118.

Chan, C., & Lam, C. K. (2010). *Conceptual change and epistemic growth through reflective assessment in computer-supported knowledge building*. Paper presented at the International Conference of Learning Sciences, Chicago.

Chan, C., & Van Aalst, J. (2003). Assessing and scaffolding knowledge building: Pedagogical knowledge building principles and electronic portfolios. In U. Hoppe, B. Wasson & S. Ludvigsen (Eds.), *Support for Collaborative Learning 2003 -Designing for change in networked learning environments*: Kluwer Academic Publishers.

Craig, J., & Cairo III, L. (2005). Assessing the relationship between questioning and understanding to improve learning and thinking (QUILT) and student achievement in mathematics: A pilot study. *Appalachia Educational Laboratory at Edvantia*.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6-28. doi: 10.1016/j.compedu.2005.04.005

Dillon, J. T. (1988). *Questioning and teaching. A manual of practice*: ERIC.
Dönmez, P., Rosé, C., Stegmann, K., Weinberger, A., & Fischer, F. (2005). *Supporting CSCL with automatic corpus analysis technology*. Paper presented at the Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!, Taipei, Taiwan.
Gweon, G., Rosé, C., Carey, R., & Zaiss, Z. (2006). *Providing support for adaptive scripting in an on-line collaborative learning environment*. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems, Montreal, Quebec, Canada.
Hakkarainen, K. (2003). Progressive inquiry in a computer‐supported biology class. *JOURNAL OF RESEARCH IN SCIENCE TEACHING, 40*(10), 1072-1088.
King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.
Kumar, R., Rosé, C., Wang, Y.-C., Joshi, M., & Robinson, A. (2007). *Tutorial Dialogue as Adaptive Collaborative Learning Support*. Paper presented at the Proceeding of the 2007 conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work, Los Angeles, California, USA.
Lee, E. Y. C., Chan, C. K. K., & van Aalst, J. (2006). Students assessing their own collaborative knowledge building. *International Journal of Computer-Supported Collaborative Learning, 1*(278-307).
Mayfield, E., & Rosé, C. (2010). *An interactive tool for supporting error analysis for text mining*. Paper presented at the Proceedings of the NAACL HLT 2010 Demonstration Session, Los Angeles, California.
Mayfield, E., & Rosé, C. (2011). *Recognizing authority in dialogue with an integer linear programming constrained model*. Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, Portland, Oregon.
Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning, 7*(2), 285-305. doi: 10.1007/s11412-012-9147-y
Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 3*(3), 237-271. doi: 10.1007/s11412-007-9034-0
Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge based questioning by children. *COGNITION AND INSTRUCTION, 9*(3), 177-199.
Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences, 3*(3), 265-283. doi: 10.1207/s15327809jls0303_3
Scardamalia, M., & Bereiter, C. (2006). Knowledge building. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). Cambridge, UK: University Press Cambridge.
Stegmann, K., Mu, J., Gehlen-Baum, V., & Fischer, F. (2011). The myth of over-scripting: Can novices be supported too much? . In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 Conference proceedings* (Vol. I - Long papers pp. 406-413). Hong Kong, China: International Society of the Learning Sciences.
Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*(1), 29-48.
van Aalst, J. (2009). Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses. *International Journal of Computer-Supported Collaborative Learning, 4*(3), 259-287. doi: 10.1007/s11412-009-9069-5
van Aalst, J., Chan, C., Tian, S. W., Teplovs, C., Chan, Y. Y., & Wan, W.-S. (2012). *The knowledge connections analyzer*. Paper presented at the The future of learning: Proceedings of the 10th international conference of the learning sciences (ICLS 2012) Sydney, Australia.
Vygotsky, L. S. (1986). *Thought and language* (2nd ed.). Cambridge, MA: MIT Press.
Zhang, J., Hong, H. Y., Scardamalia, M., Teo, C. L., & Morley, E. (2011). Sustaining knowledge building as a principle-based innovation at an elementary school. *Journal of the Learning Sciences, 20*(2), 262-307.
Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge building communities. *Journal of the Learning Sciences, 18*(1), 7-44.

Acknowledgement

The data collection for this study was supported by a grant from the Universities Grants Council of Hong Kong "Developing a Teacher Community for Classroom Innovation through Knowledge Building" (Grant HKU 740809H) and a Faculty Research Fund (Grant 201209176200).

Enacted Misconceptions: Using Embodied Interactive Simulations to Examine Emerging Understandings of Science Concepts

Robb Lindgren and Michael Tscholl, University of Illinois Urbana-Champaign,
1310 S. 6th St., Champaign, IL, 61820
Email: robblind@illinois.edu, m.tscholl@cs.ucl.ac.uk

Abstract: In this paper we describe an approach to examining and potentially diagnosing middle school students' misconceptions and emerging understanding about science concepts using immersive and interactive simulations. There are varying views in the literature on the nature of science misconceptions and the role they ought to play in learning interventions, but we focus here on their manifestation in physical activity as opposed to their detection via standardized inventories. We describe a novel framing of incorrect and emerging notions of science, and we illustrate this framing with descriptions of students' embodied interactions in an immersive digital simulation of planetary astronomy. Through live observation and video analysis we identified 9 misconceptions that were made visible through the student's bodily activity rather than through verbal accounts. We conclude with a discussion of how these same diagnostic environments may be used for instruction and remediation.

Introduction

Recent developments in interactive and immersive digital environments have opened new opportunities not only for instruction, but also for understanding and diagnosing students' misconceptions and developing ideas. Commonly thought to arise out of everyday experience in the physical world, misconceptions are qualitative intuitions or pre-conceptions that are resistant to change, even when confronted with traditional forms of instruction. While there are many viewpoints in the literature on the precise nature and origins of science misconceptions (e.g., naïve theories, knowledge fragments, etc.) the methods employed to identify student misconceptions and primitive understandings of physical phenomena has consistently relied on clinical interviews and subsequently developed paper-based inventories with schematic representations and multiple-choice questions. While we acknowledge that these instruments have provided valuable insights into facets of students' scientific reasoning, we believe they may be limited in their ability to detect the kinds of flawed conceptions and primitive ideas that are revealed not through students writing and spoken words, but through their actions in the physical world. A reliance on concept inventories for probing student misconceptions continues despite the emergence of interactive technologies that make it possible for learners to "act out" their knowledge—which includes misconceptions and emerging ideas—and make qualitative real-time predictions of the behavior of a physical system (e.g., Dede, Salzman & Loftin, 1996).

In this paper we put forth the notion of "enacted misconceptions" and we discuss the potential to use interactive technologies to make explicit the intuitive and developing ideas about science as revealed through a learner's bodily actions. This approach is based on recent advances in understanding the role of embodiment in conceptual development and learning. We explicate this framing by describing a set of specific enacted misconceptions that were revealed by middle school students using a full-body immersive simulation of how objects move in space. Finally, we discuss the potential for using the same kinds of interactive environments utilized for diagnostic purposes as a means of instruction and building up correct intuitions.

The Embodied Origins (and Expressions) of Science Misconceptions

Research on conceptual development in science over the last several decades has vigorously investigated the causes underlying students' consistent and robust errors with common science problems. They are variously described as common sense beliefs (Halloun & Hestenes, 1985), preconceptions (Clement, 1982), or as an 'intuitive sense of mechanism' (diSessa, 1993). As for the source of these intuitions or preconceptions, researcher frequently point to the influence of everyday experience in the physical world. For example, at the source of students' idea that "motion implies force" are everyday observations that objects stop moving unless a force is applied to them. The absence of an observable agent causing the counterforce (e.g., friction) likely accounts for students' difficulty in appreciating Newton's First Law.

Experiences giving rise to faulty intuitions are normally thought to be detached observations of everyday situations (e.g., watching a ball roll across the floor). Less attention has been given to the possibility that personal bodily sensations and experiences may be the basis for faulty conceptions about the physical world. For example, people traveling in a car that steers sharply will feel "pushed" against the door, potentially leading to the (incorrect) notion that centripetal force is directed outward from the center of curvature. Not only is it likely that these ideas about force *originate* from body-based actions in the world, it is also possible that these ideas are primarily *expressed* through embodied interactions. Thus, individuals in a sharply turning car

 © ISLS

may very well report consistently that the force is outward, but people looking at a diagram of circular motion are likely to vary widely in their interpretation of the system's forces (though perhaps with still very few people stating correctly that the force is directed inward). The point is that while most significant learning theories, including the Piagetian model, place prominence upon people's physical actions in the environment, knowledge is still typically inferred from whatever abstract and symbolic representation can be articulated verbally.

Contemporary research and theoretical developments in psychology and philosophy have challenged the symbolic view. Goldin-Meadow, Alibali, & Church (1993), for example, argue that children's body movements (and specifically gestures) play an important role in cognitive development and facilitating the process of learning (c.f., Abrahamson, Trninic, Gutierrez, Huth, & Lee, 2011; Antle, 2013). Theories of embodied cognition maintain that even complex concepts are derived via the perception of our bodies (Johnson, 1987). On the more radical side of embodied cognition theory, some researchers have argued that all cognition is essentially a process of activating simulations of bodily action (Barsalou, 1999).

If it is the case that cognition is inseparable from the actions and sensations of the body, then it makes sense that one would look to bodily behavior as a means of examining a person's understanding (or misunderstanding) of scientific phenomena such as principles of physics. In particular, it makes sense to create diagnostic environments grounded in the perspective of one's body (Gallagher, 2005; Lindgren, 2012). Advances in computer technology and specifically in the area of augmented and mixed reality make these kinds of immersive diagnostic environments possible.

Traditional Approaches to Examining Misconceptions

The diagnosis of students' misconceptions in physics primarily utilizes paper-based schematic representations (or web-based equivalents) often accompanied by explanatory or descriptive text. The examples describe or represent a situation requiring the application of the rules and principles of physics. For example, students may be asked to describe the forces acting on an object thrown upwards or, more often, select among 4-5 drawings representing the correct forces as arrows. Through examples of this kind, researchers and teachers gain insight into whether students hold naïve concepts of physics; in this particular case, students may select an answer containing 2 forces (one upwards and one downwards, this latter the force of gravity) while the only force acting on the object is the force of gravity. The foundational studies on physics misconceptions (McCloskey, Caramazza, & Green, 1980; Clement, 1982; Halloun & Hestenes, 1985) combined these representations with clinical interviews to elicit reasoning and explanations. From these studies, multiple-choice tests were developed based on common answers given to these questions (e.g., Force Concept Inventory (FCI); Hestenes, Wells, Swackhamer, 1992). Note, however, that even in their pre-multiple-choice origins, the misconceptions questions were rooted in students' verbal accounts of their reasoning while passively observing diagrammatic representations. We certainly do not deny that these traditional instruments have value in revealing errors in students' verbal processing of physics problems, but if one accepts the idea of misconceptions as a deeply intuitive misunderstanding expressed through behavior, then it is quite possible that these paper-based instruments are insufficient for fully characterizing student conceptions.

The impact of representational formats in students' performance on assessments is, as in a variety of other subjects, well known in physics learning. For example, Meltzer (2002) conducted research on the discrepancy between student performance with physics problems presented with sketches and text versus graphs and text. He found that answers designed to reveal student misunderstanding differed in relation to the formats used. The study suggested that students had acquired sophisticated strategies to interpret graphical representations in relation to specific questions, answering them often correctly; but interviews showed that students still did not have an intuitive understanding of physics governing the behavior of real-world objects. Hestenes (1996) emphasizes that the success of teaching of physics depends crucially on the representations being used, with multi-layered integrative models capturing the dynamic of the physical world being among the more successful ones. Likewise, success in assessments depends on the way in which knowledge is enacted, whether it is text, graphical representations or, indeed, computer-based dynamic simulations. Theoretical support of this position is provided by diSessa's (1993) notion of "knowledge in pieces." He proposes that students do not hold a systematic and consistent naïve theory of physics but rather a loose collection of notions about specific situations. This may mean that placing students in contexts in which they have not yet been tested may show a wider range and diversity of conceptions than normally uncovered by paper-based tests. Further, embedding students in action-based environments that more closely resemble the contexts in which this knowledge would be applied would presumably allow the diagnostic process to focus on the ideas and situational understandings that are most in need of further development and remediation. Indeed, part of the success of novel digital learning environments, and especially interactive desktop-based simulations (e.g., Adams et al., 2008), may derive from the ability of these systems to tap into learners' misconceptions not revealed by "static" multiple-choice questions, though clearly systematic studies are needed to investigate this possibility.

The Potential to Use Interactive Learning Technologies to Examine Misconceptions

Previous attempts to use novel and technology-enhanced learning environments to uncover incorrect ideas and intuitions about the physical world are sparse. Bates and Galloway (2010) surveyed recent uses of diagnostic methods in a mechanics courses and found that sketches depicting real-world objects combined with text and multiple choice questions remained the standard. This includes examples where simulations and other digital tools were used for learning and instruction, but when it came to assessing the presence and persistence of misconceptions, these studies still relied on traditional inventories and diagrammatic representations. This is to a large degree justified by the availability of a very large amount of data collected over decades relying on standard examples and therefore available for comparison. Though these data continue to confirm the internal validity of standardized tests (i.e., coherence of answers across test examples in relation to assumed deep-seated misconception), little has been done to examine how misconceptions might be changing or vary in their appearance given the availability of contemporary learning technologies. Unfortunately, the impact of learning technology is often seen as increasing student engagement and interest, not as an expanded context for application and assessment of critical concepts.

Unlike more passive and didactic instructional approaches, interactive learning technologies (e.g., computer simulations) focus on actions in realistic contexts, and they are designed to give learners opportunities to experience the effect of their actions (Dede, 1995; de Jong & van Joolingen, 1998; Lindgren & Schwartz, 2009). In addition to creating powerful learning interventions, these environments also permit researchers and instructors to observe students' real-time actions and choices that may reveal problematic intuitions and misconceptions about science. By placing the learner in a realistic environment, she may mobilize deep-seated intuitions, exposing them to experimentation and reflection, and making them amenable for change. It is important to note that an interactive technology environment for exposing primitive concepts does not need to be completely unstructured. Guidance through the use of the simulation or visualization technologies can still be provided, as long as there are built-in opportunities for learners to make choices and articulate predictions. As suggested earlier, we are particularly enthusiastic about the potential for revealing misconceptions with *embodied* interactive simulations, where learners not only take action within a digital environment, but they physically enact their understanding through gross gestures or full-body movement. In the next section we describe the design of one such embodied interactive environment.

The *MEteor* Simulation of Planetary Astronomy

The remainder of the paper examines a particular simulation environment and how we used it to identify several specific body-based misconceptions—what we have termed enacted misconceptions—pertaining to an understanding of how objects move in space. The *MEteor* simulation that we developed at our lab study is a large full-body simulation of planetary astronomy. Learners launch a virtual asteroid into a simulation of outer space (planets with gravitational forces, etc.), and run with the asteroid, predicting its trajectory in real time. The simulation consists of 4 different levels that require the learner to accomplish certain tasks, such as hitting a target with the asteroid that is located behind a planet, or putting their asteroid completely in orbit around a planet. The simulation was built in collaboration with physicists and physics educators to ensure accuracy of the science principles being represented.

The system is made interactive with laser-scanning technology that monitors the learner's movements across the floor display, and gives feedback on her success in following the asteroid. While the learner can move freely across the simulation, cuing mechanisms and feedback have been implemented to help guide the learner. For example, if the learner ever gets more than about 4 feet away from the asteroid it will disappear and the leaner will have to re-launch. Nonetheless, there is still ample room for learners to make inaccurate predictions and move in ways that are counter to core physics principles.

Enacted Misconceptions of How Objects Move in Space

Fifty-seven students from local middle schools used the *MEteor* full-body simulation for about 15 minutes each, working independently through the objectives of the 4 different levels. Students were given basic instructions on how to interact with the simulation but were largely free to explore and move in ways that they felt would best accomplish the objectives (i.e., hitting a target). We rely on the analysis of students' behavior as visible in video recordings as well as our live observations and field notes when the study was being conducted. In our review, we observed fairly consistent behavior indicating incorrect and emerging ideas about how objects move in space. These misconceptions and primitive ideas are much more specific than the kinds of misconceptions typically diagnosed by traditional instruments, reinforcing Liu & McIsaac's (2005) argument that the manner in which naïve physics intuitions gets applied depends on the kinds of questions, prompts, situations, and objects they are exposed to. There is good reason to believe that participants possess these incorrect ideas because they are actively trying to make accurate predictions to succeed in a game-like environment (as opposed to inferring

Table 1. Enacted misconceptions that were observed from video analysis of middle school students using the *MEteor* simulation.

***MEteor* Level and Objective**	**Observed Behavior**	**Specific Misconceptions**
Level 1: Launching the asteroid at a distant target (no other objects present in the simulation)	Participants frequently slow down and occasionally speed up when traveling to the target.	1.1. Moving objects in space behave like objects moving on Earth, which typically slow down over time
Level 2: Launching the asteroid at a target with a large planet nearby	Participants almost always begin by launching the asteroid the same way they did in Level 1. Participants also typically fall behind the asteroid, i.e. fail to predict gravitational acceleration	2.1. Gravity doesn't have a significant effect on objects that are relatively far away 2.2. Objects being pulled by planets will move with a constant speed toward the planet
Level 3: Launching the asteroid at a target with a small planet directly in front of it	Participants typically start by launching the asteroid from the far sides of the simulation platform in hopes of hitting it at an angle Participants are often lagging behind the asteroid and sometimes not anticipating strong gravitational acceleration and bending Participants frequently move as though they expect the gravity to be much stronger, responding similarly to how they moved with the larger planet in Level 2	3.1. Gravitational forces exerted by planets will move objects in motion directly towards the planet 3.2. The mass of a planet (different sized planets of equal density) does not affect the magnitude of gravitational force
Level 4: Putting the asteroid into orbit	The most consistent problem was that participants did not launch the asteroid tangentially, or slowly enough When participants successfully got the planet into orbit, they often moved around the planet with constant speed Most participants also moved as though they expected the orbits to be circular	4.1. Objects traveling directly at a planet will get into orbit by being swept up by "orbital forces" surrounding a planet 4.2. Objects in orbit move at a constant speed around the planet 4.3. Orbits are circular 4.4. Orbits require inherent force to be maintained

these misconceptions from their verbalizations). This configuration leads the participants to mobilize their ideas and intuitions and make them visible.

Table 1 lists the misconceptions that we observed in *MEteor*. These are organized according to the simulation level in which the enacted misconception was typically exhibited, and they are accompanied by the behavioral observations that suggest the primitive understanding. Figure 1 shows participants exhibiting a couple of the misconceptions described in the table. To be sure, observations of some of the students' behavior are consistent with ideas tested in traditional inventory assessments. This includes the observation that students frequently launch the asteroid with a force that is not required to get the asteroid to the target, suggesting the common misconception that objects stop moving quickly because their internal force dies out. That said, applying the notion that students have an "impetus" theory of motion to this particular simulation context is not sufficient to explain the list of enacted misconceptions in Table 1—it does not explain for example why a child



Figure 1. Left: Student exhibiting enacted misconception 2.2. The student's body is moving at constant speed, but the asteroid depicted in the simulation is accurately responding to the force of gravity by accelerating and is about to crash into the planet. Right: Student is exhibiting enacted misconception 4.2. The asteroid depicted in the simulation has just sped up as it moves closer to the planet, but the student had assumed the orbiting asteroid would move at constant speed and almost falls over in an attempt to catch up to it.

does not slow down when they are enacting an asteroid's orbit around a planet and the asteroid is at a point in the orbit that is a further distance from the planet.

Of particular value with this simulation experience were the insights it gave to how children think about phenomena that cannot easily be observed or depicted in a diagram, such as the less-local effects of gravity. Observed behavior from level 3 of the simulation demonstrated that students often underestimate the impact of gravity not only on the asteroid's speed but also on the direction of its trajectory. The way that students typically angle their launches suggests that they are not initially taking into account that gravity's influence is persistent even at far distances. Equally interesting was how students seemed to reason about orbits. Many students would use the word "ellipse" to describe an orbit, but their bodies suggested that they believed orbits to be perfectly circular. Likewise, a participant would typically move as though they anticipated the asteroid to travel at a constant speed in its orbit, expressing surprise and sometimes nearly falling over as they scramble to catch up with the asteroid accelerating when it gets closer to the planet.

Especially problematic were participant's attempts to put the asteroid into orbit in the first place. This often entailed direct launches at the planet, perhaps suggesting that they believed there was some sort of invisible orbital field that would suck up the asteroid and put it into circular motion, like an object dropped into a draining sink. The enactments we observed in several of the levels suggested difficulty understanding the relationship between an object's initial velocity and gravitational forces exerted on the object; gravitational acceleration is a phenomenon that is difficult to observe on Earth since we typically can only see objects falling for short periods of time. While it may not be surprising that middle school students were challenged by these ideas, it is important to recognize that it would not be easy to pose a question on paper or even verbally to elicit these intuitions and misconceptions.

While the focus here has been on diagnostics and identifying the kinds of misconceptions that are only revealed through embodied interactions, there is clearly the potential to utilize these same interactive environments as a means to instruct and evolve primitive conceptions. Once an enacted misconception is identified, such as the tendency to represent an orbiting object with constant speed, targeted feedback administered in real time and exposure to varying contexts that highlight critical relationships can be designed into the simulation. Importantly, the context of diagnosis is not separated from the context of remediation, as often occurs in traditional instructional contexts. The enactment of misconceptions is not a symptom of faulty knowledge, it is itself an expression of knowledge, that when exposed and made visible provides opportunities for direct intervention and development.

There is great promise in using embodied and interactive simulations to examine primitive conceptions, but there are certainly limitations to this approach that must be considered as well. There are limitations, for example, on what actions the human body can perform such that it may be difficult to express accurate and complex ideas. It would be near impossible, for instance, for *MEteor* participants to have moved fast enough to represent the velocity with which an asteroid is capable of traveling around a planet (relative to its approaching velocity). The inability to recreate this pattern of motion with one's body certainly should not be taken as an indication that they do not understand the underlying physics. It is also important not to assume that correctly

enacting a science phenomenon means that a person necessarily understands the causes of the phenomenon. Accurate performance can sometimes occur accidentally or is based on observations without appreciating underlying mechanisms. It is important, therefore, for interactive learning environments to elicit embodied performances that reflect deep understanding as opposed to surface-level descriptions of events.

Concluding Remarks

Recent models of cognition and learning have emphasized its embodied nature understood as the bodily enactment of thinking (Gallagher, 2005; Wilson, 2002). These views contend that there is a deep connection between thinking and bodily activity, and prescribed forms of physical movement may aid cognitive processes and increase the propensity for learning (Goldin-Meadow, 2005). In recent years increasing attention has been given to instructional environments and methods that incorporate body movements with the aim to help people build connections between their physical experiences and important principles and ideas in STEM domains (Abrahamson & Lindgren, in press; Goldin-Meadow, Cook, & Mitchell, 2009). Embodied learning is made especially relevant in relation to the development of immersive interactive environments that allow seeding of bodily activities (Lindgren & Johnson-Glenberg, 2013). However, while these systems show considerable promise to develop new forms of learning, less attention has been given to the possibility of using these systems to identify primitive and emerging conceptions. While the important role of interactivity in learning has being recognized and the use of interactive systems is being utilized in a range of STEM learning environments, often the impact of interactivity is conceptualized in relation to increased engagement, enjoyment, motivation and other indirect measures of learning. Now that immersive and interactive digital environments can confront learners with novel and highly contextualized scientific events, the possibility for access to different forms of knowledge than what is available in traditional diagnostics needs to be given serious consideration.

References

Abrahamson, D., & Lindgren, R. (in press). Embodiment and embodied design. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd edition). Cambridge, MA: Cambridge University Press.

Abrahamson, D., Trninic, D., Gutiérrez, J. F., Huth, J., & Lee, R. G. (2011). Hooks and shifts: a dialectical study of mediated discovery. *Technology, Knowledge and Learning, 16*(1), 55-85.

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman, C. E. (2008). A study of educational simulations part I: Engagement and learning. *Journal of Interactive Learning Research, 19*(3), 397-419.

Antle, A. N. (2013). Exploring how children use their hands to think: An embodied interactional analysis. *Behaviour & Information Technology, 32*(9), 1-17.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-660.

Bates, S.P., & Galloway, R., (2010). Diagnostic tests for the physical sciences: A brief review. *New Directions in the Teaching of Physical Sciences, 6*, 10-20.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Dede, C. (1995). The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. *Educational technology, 35*(5), 46-52.

Dede, C. J., Salzman, M., & Loftin, R. B. (1996). The development of a virtual world for learning Newtonian mechanics. In *Multimedia, Hypermedia, and Virtual Reality Models, Systems, and Applications* (pp. 87-106). Springer Berlin Heidelberg.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research, 68*(2), 179-201.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition & Instruction, 10*(2-3), 105-225.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Clarendon Press.

Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Harvard University Press.

Goldin-Meadow, S., Alibali, M.W., & Church, R.B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*(2), 279-297.

Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics, 53*(11), 1056-1065.

Hestenes, D. (1997). Modeling methodology for physics teachers. In *AIP Conference Proceedings* (Vol. 399, p. 935-955).

Hestenes, D., Wells, M., Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141-158. Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.

Lindgren, R. (2012). Generating a learning stance through perspective-taking in a virtual environment, *Computers in Human Behavior, 28*(4), 1130-1139.

Lindgren, R., & Johnson-Glenberg, M. C. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 45-52.

Lindgren, R., & Schwartz, D. L. (2009). Spatial learning and computer simulations in science. *International Journal of Science Education, 31*(3), 419-438.

Liu, X., & MacIsaac, D. (2005). An investigation of factors affecting the degree of naïve impetus theory application. *Journal of Science Education and Technology, 14*(1), 101-116.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science, 210*(4474), 1139-1141.

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics, 70*, 1259.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636.

Friendship, Participation, and Site Design in Interest-Driven Learning among Early Adolescents

Ashley Cartun, Ben Kirshner, Emily Price, Adam York, University of Colorado Boulder
Ashley.Cartun@colorado.edu, Ben.Kirshner@colorado.edu, Emily.Price@colorado.edu,
Adam.J.York@colorado.edu (1)

Abstract: Increasingly, learning scientists are recognizing the importance of studying and analyzing learning across the multiple settings of youths' lives. We hypothesize that the potential for positive long-term outcomes for youth in interest-powered learning environments is shaped by the degree to which programs cultivate personal connections that can expand access and strengthen participation in settings rich with resources for interest development. To investigate this hypothesis, we draw on evidence collected by youth researchers as part of a study of five learning environments that aim to support cross-setting pursuit of interests. Findings from this study support claims that youth do facilitate access to valued learning spaces for their peers, and that friendships can be central to sustaining interest in activities. The data from this youth ethnographic study supports placing a priority on *relationships*, as well as content, within sites for interest driven learning.

Introduction

Increasingly, learning scientists are recognizing the importance of studying and analyzing learning across the multiple settings of youths' lives (Bevan, Bell, Stevens, & Rafzar, 2012; Jackson, 2011). One of the key catalysts for cross-setting learning is *interest* (Barron, 2006; Bell, Bricker, Tzou, & Baines, 2012), which refers to engagement with particular content that develops through interactions with others around that content (Hidi & Renninger, 2006). Interest-driven learning environments are ones that seek to organize activities that allow youth to pursue and develop existing interests and develop new ones (Azevedo, 2013; Edelson & Joseph, 2004; Ito et al., 2013).

In this paper, we explore the role of peers and friends in shaping interest development and sustaining participation in interest-driven learning environments. We hypothesize that the potential for positive long-term outcomes for youth in interest-powered learning environments is shaped by the degree to which programs cultivate personal connections that can expand access and strengthen participation in settings rich with resources supportive of interest development. To investigate this hypothesis, we draw on evidence collected by youth researchers as part of a study of five learning environments that aim to support cross-setting pursuit of interests. Our study findings point to exciting possibilities, such as how to design for learning spaces that support friendship as one path to expanding possible futures for youth.

Theorizing the Role of Peers in Participation in Interest-Driven Learning Across Settings

Interest-driven and out of school settings are particularly valuable avenues for experiencing competence and leveraging expertise from their families and communities (Calabrese Barton, Tan, & Rivet, 2008; Gutiérrez, Morales, & Martinez, 2009). Increasingly, policy makers and researchers alike have called for tighter linkages between formal and informal settings for learning, to better leverage diverse youths' expertise and enable deeper forms of "life-wide" learning (Banks et al., 2007). Tighter linkages are especially important for providing youth with recognition for accomplishments outside of school in ways that are consequential for youth's social futures (Riconscente, Kamarainen, & Honey, 2013).

Within this line of research on cross-setting and interest-driven learning, however, there has been limited attention to the role peers play in interest-driven learning within and across settings. Peer connections are potentially important, however, because it is through networks that many young people discover new interests, and friendships developed within interest-driven learning activities may help to sustain participation in them and, subsequently, facilitate deepening of interests. Evidence from developmental psychology supports the claim that peers are an important reason why adolescents sustain engagement in sports and arts activities (Patrick et al., 1999). In some cases, peer connections may also benefit youth, because they broker access to activities where youth can develop new knowledge and skill (Dika & Singh, 2002; Stanton-Salazar & Dornbusch, 1995).

Our focus in this analysis on peer relationships is supported by evidence from studies on youth development. One reason to study peers in the context of interest driven learning sites is that many sites of this kind cater to adolescents, and during the adolescent years peers become a more important influence than in other stages of development (Berndt, 1982; Hartup, 1993; Savin-Williams and Berndt, 1990). During adolescence, as peer groups become more salient, other relationships can be redefined for young people, who often report that

they can most “be themselves" when they are with their peers (Savin-Williams and Berndt, 1990). Studies also show that peers play an important role in shaping motivation for academics (Juvonen and Wentzel, 1996), an outcome of concern when discussing cross-setting learning and the impact of interest-powered sites.

By attending to youths’ own reports of the role of peers in their experience of interest-driven learning we can begin to consider the features of learning environments that may support or inhibit development of relationships. Deepening this understanding may hold lessons for site design practices including the general point that to design appealing and effective learning spaces for young people involves attending to both content learning and peer relationship development.

Methodology

Data for this analysis was gathered from a participatory youth ethnographic study of sites where young people engage in interest-driven learning. The current study is principally an observational study, where youth are part of designed learning environments. The team developed a youth participatory research component that provided training in interviewing techniques, use of structured protocols for conducting peer and mentor interviews, and activities in which youth traced their own experiences of connected learning across setting and time. The youth ethnography took place as a series of online video conferences that were facilitated by the research team. This participatory research connected nodes of youth researchers that were geographically distributed and met in web-mediated settings (in this case Google hangouts). The infrastructure we developed, including training materials, a wiki, hangouts, email communication, and stipends, sustained research activities over a 10-week program.

This research was conducted under the umbrella of the Connected Learning Research Network (CLRN), which is investigating and analyzing the impacts of today's changing media ecology on learning (Ito et. al., 2013). This participatory youth ethnography is part of a larger mixed-methods research study that examines children’s participation in connected learning environments and the relationship of participation to valued outcomes of interest development, persistence, and school belonging.

Youth ethnography was selected as a methodology because it creates an opportunity for young people to become involved in documenting and analyzing their experiences in these environments including the kinds of opportunities available to them, and to identify what counts as a “quality” opportunity to young people (Rubin & Jones, 2007). The rationales for including participatory research in an ethnographic study of learning environments include the awareness that young people have a unique insider perspective on the activity in those spaces and can contribute in valuable ways to the construction of knowledge about what supports learning (Cammarota & Fine, 2008, Sabo-Flores, 2008). Engaging in participatory research can also serve as a positive developmental opportunity for young people in that it invites them to participate in practices of investigation and analysis of systems (Mitra, 2004). Research of this nature disrupts the dichotomy between researcher and research participants, with an expanded belief about who should generate knowledge to inform policies and practices (Cook-Sather, 2002, Morrell, 2008).

Participants

Sites

Five sites participated in the youth ethnography. Table 1 lists program pseudonyms and a description of each of the site’s central activities. The sites were geographically distributed, with one site on the west coast, one in the rocky-mountain west, one in the southeast, and two in the northeast United States.

Table 1: List of Sites

SPARK	A youth program for documentary filmmaking.
Community Bridge	A school based school community service program connected to a national initiative focused on the improvement of teaching writing.
The SPOT	An after-school youth program that offers arts and humanities courses.
NEXT School	A school where students take courses in game design, critical research and entrepreneurship.
Freespace Library	A library teen program that was in the process of designing a new technologically rich space.

Youth
Roughly five youth from each of the five sites were hired to participate in a paid internship within this research project. Most of the youth were from nondominant backgrounds. The participants included black, Latino, and white youth from low and middle income communities.

Sources of Evidence
Youth researchers recorded a total of 82 peer and 12 mentor interviews, 27 "interest stories" detailing the development of an interest or hobby, and they also submitted 22 brief digital self-introductions. Eleven youth submitted maps showing their daily and weekly routines that were accompanied by audio explanations of what they depicted on their maps.

Peer Interviews
The primary source of data used for this analysis is a set of 82 semi-structured interviews where youth at the sites recruited peers for the study. The youth researchers adapted a protocol for eliciting details about long-term, interest-driven projects, the development of expertise in an area of interest to them, and the formation of new social ties through participation in connected learning.

Mentor Interviews
Youth researchers recruited and interviewed mentors who played a role in their own interest development or learning at their site. The interviews included questions about the mentor's philosophy towards work with young people and the supports and challenges to their work.

Interest Maps
Some of the youth participants documented their everyday experience by creating a hand drawn map showing their daily and weekly routines related to an interest. These were accompanied by audio explanations of what they depicted on their maps. Elements in the map included people, places, and resources that were important to helping them pursue an interest.

Interest Stories
To deepen our understanding of interest development, we asked youth ethnographers to articulate stories of their own interests. The youth participants drew a storyboard depicting the formation of an interest, how the interest figures into their life currently, and where they see themselves taking the interest in the future. These storyboards were also accompanied by audio narrations.

Personal Introductions
At the outset of the project youth recorded brief audio introductions describing thoughts and questions relating to participating in a study of their site, and detailing some of their personal interests.

Approach to Analysis
In an effort to attain a valid coding scheme and inter-rater reliability, the research team established a multi-step coding process, which included the development, testing and refinement of codes throughout a series of coding summits. Coding clusters were developed from the project's various qualitative data sources. Our analytical process resulted in three broad groupings: one pertaining to the lead interest-based activity, another pertaining to the building of cultural capital (interest, expertise, and identity development), and a third pertaining to the building of social capital (connections with others), which also contained a series of child codes to be used for co-occurring patterns and other forms of qualitative and quantitative analysis. To test this coding scheme, third party validity testing was used, and we achieved an inter-coder agreement of at least 80% for all codes. Once validity was established, codes were refined through a series of reliability testing using kappa statistics. This multi-step coding process using defining guidelines and principles established inter-rater reliability and aligned with our research questions and theoretical framework related to connected learning practices.

Results
In our data we observed peers mentioned often in interviews and data that the youth ethnographers collected about their interest pursuits. However, peers fill different roles in people's experiences across sites. The central theme that that we highlight in our findings is that sites vary with respect to how friendship figures into participation at the site. Interviews from The SPOT contained many examples of young people being introduced to the site by peers, both friends and near-age family members. At two of the sites, Community Bridge and SPARK, the primary story emerging from interviews was that pursuit of the activities at the site led to new friendships. These two sets of results illustrate how peers might impact interest development in different, but equally important ways.

Peers as Brokers

The data from The SPOT suggests that friends and family in some cases figure prominently in how people initiated their involvement with the site. A common them in interview responses was that friends and family verbally shared information about the types of activities and opportunities for learning at The SPOT. Peers shared that this interest-powered site was "open" and "fun," and in some cases young people joined their peers to visit The SPOT just to try it. After young people started going to the site, they report that they began to develop or expand upon interests in activities at The SPOT such as making music, dancing and video production. This is an encouraging finding in part because it shows that peers in some cases do facilitate access to expansive opportunities for interest development.

One of the young people at The SPOT clearly recalls how they were introduced to the site, first from interest sparked by a friend discussing the site, and then how they found interesting opportunities that they began to pursue. "Well, I have a friend who goes there...and he told me about how he does his own music there. I got interested in it, and I asked him if I could go with him one day. He took me, and I enjoyed it so much that I signed up right away. I started doin' dance and music production myself." This type of story was common at this site, with 9 of the 20 peer interviews containing similar comments about the role of peers in introducing the young person to new activities at The SPOT, which has deepened some personal interests.

Data from The SPOT suggests that peer word of mouth can be a powerful mechanism for promoting access to valued spaces for interest development. In this case peer recruiting was an organic process, initiated because young people wanted others to know about the interesting opportunities at the site. This process was supported both by the fact that young people found the site to be offering valuable opportunities, and because the site is open to newcomers who can begin taking part in activities simply by showing up.

Forming New Connections

Many of the youth interviewed at the Community Bridge and SPARK reported that the most interesting part of being involved in this program was meeting new people. We saw this theme echoed in 20 of the 39 peer interviews across these two sites. As one interviewee stated, "...connecting with people I would say that's the most interesting thing about this project because you get to see the world through their eyes, you get to relate to them, and um you get to see stuff and do stuff that you probably haven't even thought of." At both SPARK and Community Bridge youth engaged in work that prompted them to learn skills with digital media, yet making new personal connections was more often reported to be the *most* interesting aspect of participating in the programming.

Both of these sites offered programming that revolved around joint activity. In the case of Community Bridge, students collaborated on community service oriented senior projects where they created digital stories, videos, and blogs that chronicled their work. At SPARK the youth participants worked alongside college students producing documentary films on a variety of social issue topics, most recently education reform. Through these projects team members at both sites had to find ways to work together drawing on the expertise distributed throughout their groups in order to accomplish a shared task. Also of note was that the projects at both of these sites required that participants have contact with community members outside of their organization in order to gather information to inform their projects.

Youth from these sites reported on new connections with peers who were part of their project teams, as well as new connections they formed through their community outreach. We heard similar results, though in smaller numbers, from other sites, including Freespace Library and The SPOT, where meeting new people was discussed as one positive aspect of participation.

Discussion

The findings from this youth ethnography support the assertions of other studies that emphasize the importance of peer connections in adolescence (Savin-Williams and Berndt, 1990). Peers are mentioned often as young people talk about their experiences at sites where they pursue their interests. What we find thought-provoking is that across the sites that we collected data with, there was some variation in how peers fit into the stories young people told about what brought them to sites or sustains their interests. Trying to understand the multiple forms of peer connections that young people experience in more nuanced ways will be an ongoing task of this longitudinal study. This data set supports some minor but meaningful claims about the role of peers and friendship in interest driven learning.

One claim is that youth do facilitate access to valued learning spaces for their peers. Our results foregrounded young people as brokers to valued spaces, especially at one site within our study where peer introductions was the primary mechanism through which people were introduced to the site. At this site open access supported the process of peer recruitment by allowing for young people without prior experience to come to the space through the invitation of friends or family. Unfortunately, this may not be a common experience for youth with limited access to resource rich spaces. We can think of at least two ways to expand upon this finding in future research. First, by studying the powerful appeal of peer recommendations and what encourages young

people to take up an invitation to try a new activity. Second, by investigating in closer detail the features of learning spaces that can support young people in feeling comfortable and prepared to share their experiences with peers. Both of these lines of inquiry make the peer connections central as a potential form of access to valuable resources.

Our second claim is that developing friendships is often a key element of interest-driven learning. Many times, even in the context of new media and technology rich spaces, it was personal connections that emerged as the "most interesting" aspect of activity. We suggest some consideration for research and design work in interest-driven learning spaces from this finding. The design of sites where "making new friends" was reported often centered activity around participating in a joint endeavor that required some form of outreach, such as community based research to inform a documentary film. This design encouraged both meeting new people and shared activity that promoted deep and sustained personal engagement. Friendships and engagement with activity in these cases were not detached: rather they were both integral to the positive experience that young people reported at their sites. Activity supported the development of new peer connections, and these new friendships made the sites an interesting place to engage in activity.

The data from this youth ethnographic study supports placing a priority on *relationships*, as well as content, within sites for interest driven learning. In the stories that young people shared through interviews and self-documentation friendship was a constant presence as they described their initial discovery and pursuit of interests. These findings may be of use in future research and design of interest-driven sites for learning by attending to factors within and across contexts that support peer connections.

Endnotes

(1) Authors listed alphabetically to indicate equal contribution.

References

Azevedo, F. S. (2013). The tailored practice of hobbies and its implication for the design of interest-driven learning environments. *The Journal of the Learning Sciences, 22*(3), 462-510.

Banks, J. A., Au, K. H., Ball, A. F., Bell, P., Gordon, E. W., Gutierrez, K., Heath, S. B., Lee, C. D., Lee, Y., Mahiri, J., Nasir, N. S., Valdes, G., & Zhou, M. (2007). Learning in and out of school in diverse environments: Life-long, life-wide, life-deep. Seattle, WA: UW Center for Multicultural Education

Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development, 49*, 193-224.

Barton, A. C., Tan, E., & Rivet, A. (2008). Creating hybrid spaces for engaging school science among urban middle school girls. American Educational Research Journal, 45(1), 68-103.

Bell, P., Tzou, C., Bricker, L. A., & Baines, A. D. (2012). Learning in diversities of structures of social practice: Accounting for how, why, and where people learn science. *Human Development, 55*, 269-284.

Berndt, T. J. (1982). The features and effects of friendship in early adolescence. Child development, 1447-1460.

Bevan, B., Bell, P., Stevens, R., & Razfar, A. (Eds.). (2012). *LOST Opportunities: Learning in out-of-school time*. New York, NY: Springer.

Cammarota, J., & Fine, M. (Eds.). (2008). Revolutionizing education: Youth participatory action research in motion. Routledge.

Cook-Sather, A. (2002). Authorizing students' perspectives: Toward trust, dialogue, and change in education. Educational researcher, 31(4), 3-14.

Dika, S. L., & Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. Review of educational research, 72(1), 31-60.

Edelson, D. C., & Joseph, D. (2004). The interest-driven learning design framework: Motivating learning through usefulness. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon & F. Herrera (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences* (pp. 166-173). Mahwah, NJ: Erlbaum.

Gutiérrez, K. D., Morales, P. Z., & Martinez, D. C. (2009). Re-mediating literacy: Culture, difference, and learning for students from nondominant communities. Review of Research in Education, 33(1), 212-245.

Ito, M., Gutiérrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K. & Watkins, S. C. (2013). Connected learning: an agenda for research and design. Digital Media and Learning Research Hub.

Hartup, W. W. (1993). Adolescents and their friends. New Directions for Child and Adolescent Development, 1993(60), 3-22.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127.

Jackson, K. (2011). Approaching participation in school-based mathematics as a cross-setting phenomenon. *Journal of the Learning Sciences, 20*(1), 111-150.

Juvonen, J., & Wentzel, K. R. (Eds.). (1996). Social motivation: Understanding children's school adjustment. Cambridge University Press.
Mitra, D. (2004). The significance of students: can increasing" student voice" in schools lead to gains in youth development?. The Teachers College Record, 106(4), 651-688.
Morrell, E. (2008). Critical literacy and urban youth: Pedagogies of access, dissent, and liberation. New York: Routledge.
Patrick, H., Ryan, A. M., Alfred-Liro, C., Fredricks, J. A., Hruda, L. Z., & Eccles, J. S. (1999). Adolescents' commitment to developing talent: The role of peers in continuing motivation for sports and the arts. *Journal of Youth and Adolescence, 28*(6), 741-763.
Riconscente, M. M., Kamarainen, A., & Honey, M. (2013). STEM badges: Current terrain and the road ahead. New York, NY: New York Hall of Science.
Rubin, B. C., & Jones, M. (2007). Student action research: Reaping the benefits for students and school leaders. NASSP Bulletin, 91(4), 363-378.
Flores, K. S. (2007). Youth participatory evaluation: Strategies for engaging young people (Vol. 14). John Wiley & Sons.
Savin-Williams, R. C., & Berndt, T. J. (1990). Friendships and peer relations during adolescence. In S. S. Feldman & G. Elliott (Eds.), At the threshold: The developing adolescent (pp. 277-307). Cambridge, MA: Harvard University Press.
Stanton-Salazar, R. D., & Dornbusch, S. M. (1995). Social capital and the reproduction of inequality: Information networks among Mexican-origin high school students. Sociology of education, 116-135.

Using Models for Reasoning and Content Learning: Patterns of Bootstrapping Towards Earth Science Understandings

Ann E. Rivet, Cheryl A. Lyons, Alison R. Miller
Teachers College Columbia University, 525 W. 120th Street, New York, NY 10027
rivet@tc.columbia.edu, cal2154@tc.columbia.edu, mar2218@tc.columbia.edu

Abstract: A key aspect of using scientific models and other representations as cognitive learning tools is the reciprocal relationship between understanding the nature of models as representations, and understanding the specific concepts and phenomena that the model is intended to represent. However, challenges exist regarding how to describe and measure indicators of this reciprocity. We explored the ways in which 8th and 9th grade students utilized physical dynamic tabletop models towards developing sophisticated understandings of full scale Earth System processes. This approach allowed us to identify and describe evidence of the "bootstrapping" that occurs between understanding the model as a scientific representation, and understanding the science concepts of the represented entities, configurations, motions, and emergent phenomena in the real Earth System. We argue that this notion of bootstrapping is a productive means to conceptualize and support the development of students' epistemological understandings of both scientific models and the represented science concepts.

Introduction

One of the challenges with the current discourse in science education around scientific practices, and modeling practice in particular, is how the development of these practices interplays with the development of sophisticated target content understandings. A key aspect of using scientific models and other representations as cognitive learning tools is the reciprocal relationship between understanding the nature of models as representations, and understanding the specific concepts and phenomena that the model is intended to represent (Schwarz et al., 2009). Although the literature fully acknowledges the intertwined relationship between the two, specific attempts to conceptualize, describe, and measure how these different but related constructs evolve across students' learning continuum are still in their infancy. In particular, challenges exist regarding both how to describe and measure indicators of this reciprocity, and the conditions under which the interplay between students' understanding the nature of models as representations (a key aspect of modeling practice) and robust conceptual understanding developed through working with such models is most productive for learning.

Our work attempts to frame this important intersection of practice and conceptual learning. Specifically, we sought to address the following question: what is the nature of the relationship between students' demonstrated content understandings and the sophistication of their analogical reasoning around representative models? We articulated a progressive analogical reasoning construct to describe the ways in which students develop in their ability to conceptualize more abstract and generalized understandings of both models as representations, as well as concepts and phenomena that are represented in specific models. Through in-depth interviews and written assessments, we explored the ways in which 8th and 9th grade students utilized physical dynamic tabletop models towards developing sophisticated understandings of full scale Earth System processes. This approach allowed us to identify and describe evidence of the "bootstrapping" that occurs between understanding the model *as* a scientific representation, and understanding the science concepts *of* the represented entities, configurations, motions, and emergent phenomena in the real Earth System (Carey, 2004; Kurtz, Mao & Gentner, 2001). We argue that this notion of bootstrapping is a productive means to conceptualize the development of students' epistemological understandings of both scientific models and the represented science concepts, and should be further explored and supported through instructional approaches and other cognitive tools embedded in science learning experiences across the K-12 continuum.

Conceptual Framework

Challenges Particular to Earth Science Learning

Science is about developing understandings and explanations of phenomena of the natural world. To the greatest extent possible, much of science education involves giving students direct experiences with such phenomena. Yet it is not possible to bring many of the important phenomena of Earth Science into the classroom setting for students to explore. Key Earth system processes, such as eclipses, ocean currents, and differential heating of the atmosphere, are beyond a students' tangible grasp. The most common way to address this challenge is to make extensive use of a wide array of representation types across the Earth Science curriculum (Kastens & Rivet, 2010), including conceptual models. Conceptual models are defined by the

National Research Council (NRC, 2012) as diagrams, physical replicas, mathematical representations, analogies and computer simulations that are simplified structural, functional, or behavioral analogs for the phenomena being represented, and can be used to generate explanations and predictions. Specifically, we focus our research on dynamic tabletop models that change or move. Research has found that such models are engaging for students, and have the ability to mirror the use of models in authentic science practice to both represent and develop new understandings (Neressian et al., 2003). However, there are known challenges with using such models in classrooms, including documented cognitive leaps of scale and rate, and instructional approaches that are often focused on the details of the model rather than on students' use of the model to develop understandings of the represented Earth System. To address these challenges in support of Earth Science learning, there is a need for greater understanding of how exactly students interpret and reason with physical models, and what kinds of supports (from the teacher, instructional materials, and the model itself) are most effective in guiding students' use of such models towards deep understandings of the Earth.

Modeling Practice

There is a distinction made in the literature between models (particularly conceptual models) and modeling practice. The practice of modeling, as described by Schwarz et al. (2009), is a weaving together of both the active engagement with the elements of modeling, and the understanding of the rationale and norms that guide the practice, referred to as *meta-modeling knowledge* (Schwarz & White, 2005). Meta-modeling knowledge includes understandings of how models are used, why they are used, and what their strengths and limitations are. Thus simply working with physical representations is claimed to be insufficient for students to develop an understanding and appreciation of modeling practice. Rather, it is through this combination of engagement with and knowledge of modeling, that students develop a more robust sense of how science works and the nature of the knowledge that science produces (Schwarz et al., 2009).

One of the persistent questions around the conceptualization of students' development of science practices in general, including modeling practice, is the nature of the reciprocal relationship between developing understandings of the practice itself and understandings of the specific scientific concepts engaged through the practices. Researchers have argued that models and the real world phenomena that they represent exist as a dialogic: it is through analyzing phenomena one can glean insights into the potential elements, relations, operations, and rules that govern and constrain the model; while concurrently, the model allows for the generation of new explanations and predictions regarding the targeted phenomena (Schwartz et al., 2009). The National Research Council (2012) goes further to state that developing an understanding of models and their role in science can help learners construct and revise their own mental models of phenomena, which in turn results in more robust reasoning and a deeper understanding of science concepts. However, as strong as this claim is in theory, there is scant evidence to illustrate such reciprocity in actual student learning. Due in part to measurement challenges, science practices such as modeling are often examined and evaluated in the abstract, apart from the disciplinary content focus in which the modeling practice is embedded. Therefore the question still remains regarding the nature of the conceptual and epistemic science learning and meaning making that is gained through engagement with modeling practices around targeted concepts and phenomena under study.

Learning from Models: Analogical Reasoning

In light of these challenges we were interested in exploring further the nature of how students come to understand the 'representation-ness' of models, and how the models are understood to serve as analogies for phenomena that are too big, too slow, or too intangible to be observed directly. To shape our thinking, we drew heavily from the literature on analogical reasoning, and in particular the work of Dedre Gentner. Gentner's structure mapping framework for analogy (e.g., Gentner, 1983) focuses the process of establishing a structural alignment between a familiar source (in this case, a physical model in front of students) and an unfamiliar target (such as a large-scale Earth process like atmospheric circulation or subduction at plate boundaries). Gentner's framework distinguishes among different forms of similarity that may exist between the source and the target, and articulates a set of implicit rules for mapping knowledge about the source onto the target. This and other educational research demonstrates how the power of analogy comes from the relationships between objects rather than from the attributes of objects themselves, and that the most powerful analogy-derived insights come from the existence of higher-order relations such as causality that correspond between the source and the target.

Building from Wilson's (2005) approach to construct modeling, in our work we identified three key levels of analogical reasoning regarding the correspondences and non-correspondences between models and the Earth System that frame an increasingly sophisticated way that students may come to use models to develop robust understandings of Earth Science concepts and phenomena (see Figure 1). To illustrate these three levels, we describe the reasoning that students may engage in around an exemplar model of the phases of the moon (Figure 2). In this model, a basketball is placed on a stand in the front of the room, with a small plastic doll taped to a point about half-way between the top and the mid-line, oriented so it is facing the classroom. A bright light is placed to the side of the basketball. The instructor then moves a smaller yellow lacrosse ball around the

basketball, at a sufficient angle so that the light from the lamp continually illuminates one side of the yellow ball as it moves around the basketball. Using this model, we describe the three levels of reasoning about correspondences and non-correspondences shown in Figure 1 that we believe users of this model would engage in while coming to better understand the target phenomena, that of the observed phases of the Moon from Earth.

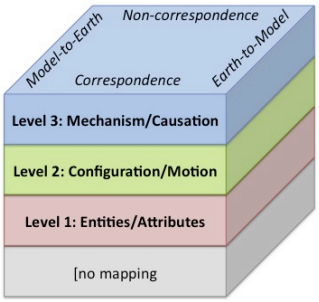


Figure 1: Levels of students' analogical reasoning between a physical model and the Earth System



Figure 2: Model of the phases of the Moon

The first, and most basic, level of analogic mapping has to do with identifying and understanding the correspondences and non-correspondences between entities in the model and entities in the Earth System. By 'entities' we are referring to both specific object mapping and mapping of the characteristics of those objects. In the example, reasoning about correspondences and non-correspondences at the entity level would include identifying and naming the basketball as representing the Earth, the yellow lacrosse ball as representing the Moon, the bright lamp as representing the Sun, and the doll as representing an observer on Earth. Likewise, non-correspondences at the entity level would include noting that the Earth is not actually orange like the basketball, that the lamp is much smaller than the real Sun, and that the stand holding the basketball has no correspondence to any entity in the real Earth System.

The second level of reasoning that users of this model would engage in involves considering either the configuration or arrangement of entities with respect to each other in the model as corresponding to similar configurations of entities in the Earth System, or the motion of entities with respect to other objects in the model as corresponding to similar motions of objects in the Earth System. For example, students may begin to reason that the motion of the yellow lacrosse ball around the basketball may correspond to the motion of the moon around the Earth. Non-correspondences of configuration or motion would also be recognized, including the fact that the motion of the moon in orbit around the Earth is considerably slower than the motion of the yellow ball in the model. Similarly, the basketball is not rotating in the model, whereas in the Earth System the Earth would be rotating concurrently to the moon orbiting around the planet.

The third and most sophisticated level of reasoning involved recognizing the phenomena as it emerges or develops in the model and identifying the cause or mechanism which drives that phenomena, and mapping that mechanism or cause such that it corresponds to the same mechanism or cause in the real Earth System. In the example, students would recognize that the doll on the basketball was observing the illumination of the yellow ball change as it orbited the basketball, resulting in 'phases' of the yellow ball at different times. This emergent phenomena of 'phases' corresponds to the phases of the moon that we observe on Earth, and the mechanism causing those phases in the model corresponds to the similar mechanism in the Earth System related to the orbit of the moon around the Earth with respect to the position of the Sun.

The dimensions of the construct include not only the vertical additive levels of mapping with increasing sophistication, but also the additional dimension of mapping both correspondences and non-correspondences at each level. As every representation is by definition some form of simplification or abstraction of actual real world phenomena, we believe fluency in reasoning around models includes both a conceptualization of a model's similarities and differences to the phenomena being represented. It is important for students to recognize not only the affordances but also the limitations of models, as the nature of the inherent approximations and assumptions of each model limits its range of validity and precision of predictive power (NRC, 2012). Thus in our construct, as well as throughout our assessments of students' reasoning, analogical mapping of both correspondences and non-correspondences played a prominent role.

Bootstrapping Between Modeling Practices and Analogical Reasoning

Given the conceptualized reciprocity between modeling practice and analogical reasoning around models, it then becomes important to consider the implications of this claim for describing and identifying observable characteristics of this relationship in students' science learning. We take up the notion of "bootstrapping" as a productive lens to consider the nature of this relationship. Bootstrapping is a frequent metaphor used in the literature to refer to the process of "using theory to constrain data and using data in turn to constrain, refine, and

elaborate theory" (Koslowski, 1996, p.281). It builds from the uniquely human capacity for learning and using representative symbols and relations between them, and the ability to integrate across distinctly different representational systems (Carey, 2004). The metaphor is thus a way to help explain how a learner is able to achieve conceptual endpoints that far transcend where she is starting from, particularly through the creation of successively new and more powerful mental representations. This process of conceptual bootstrapping is not additive, in the sense that the development along any single dimension or representational system does not logically follow a successively cumulative linear fashion. Rather, progress within such a reciprocal relationship is marked by co-concurrent development across two or more systems, with the sophistication of intermediate steps in each progression exceeding what would be anticipated if addressed in isolation.

Bootstrapping in our case refers to the ways in which students recognize and make meaning of similarities and differences between the physical model as a representation and their understanding of the full scale Earth System, in order to generalize and abstract to broader science principles. This perspective also promotes the perspective of conceptual models as cognitive tools (Brown, Collins & Duguid, 1989). For example, the more a learner are able to recognize and use the "tool" (conceptual model) of convection to explain both real world phenomena and tangible representations, the better she is able to understand both what the concept of convection stands for in the abstract, as well as the nature of an increasing array of instances and phenomena where the concept of convection is an appropriate explanation. The iterative movement between partial insights gleaned through the consideration of the "representation-ness" of physical models and the particular analogical similarities and differences between models and real Earth phenomena results in the development of more sophisticated models that account for previously unrepresented structures or behaviors, and thus enhanced conceptual understanding (Nersessian & Chandrasekharan, 2009).

However, evidence of this bootstrapping between models and conceptual scientific understanding is still minimally described in the literature. Missing are robust accounts of what students look like as they are productively (and not so productively) engaged in this process of learning. Such characterizations are needed in order to inform instructional strategies aimed to support students' effective engagement with and use of models across phenomena, as well as the nature of assessment tasks to evaluate and inform modeling practice and robust conceptual understandings. Our work aims to address this need by exploring the following research question: What is the nature of the relationship between students' demonstrated conceptual understanding and the sophistication of their analogical reasoning around physical models of full-scale Earth System phenomena?

Methods

Setting and data sources

The two-year study examining the nature of students' analogical reasoning around Earth System models was conducted in partnership with three 8th and three 9th grade Earth Science teachers in schools outside of a large city in the Northeastern US, which reflected a range of demographics and achievement levels. In Year 1 of the study, teacher used models as part of their typical Earth Science instruction. In Year 2, they incorporated specific instructional strategies to support students' reasoning around models (see Rivet, et al. 2013).

We developed parallel assessment activities around three Earth Science topics: phases of the moon, the cause of the seasons, and differential sorting in depositional environments. Teachers addressed each of these topics at various times across the year in lessons spanning 1-4 class periods, utilizing their own selected array of models and other representations in each lesson. For our assessment activities, we featured a researcher-run 3D dynamic model of the Earth System process that we developed, set up in the front of the room for students to observe (e.g., see Figure 2). A pre/post written assessment consisting of 20-22 short answer and multiple-choice items was developed for each topic. Items were crafted to elicit students' understanding of either a correspondence or non-correspondence between the assessment activity model and the targeted aspect of the Earth System, at one or more levels of the analogical reasoning construct. Additionally, individual videotaped interviews with selected target students from each teacher were conducted after each posttest administration to further elicit more detailed explications of their reasoning and content understanding. These interviews involved asking students to elaborate on their understanding and reasoning around a selected group of 6-8 posttest items for each topic. During these interviews, students were provided with a miniature version of the model used for the written assessments and told that they could use the model at any time to help them explain or figure out an answer. Further information about the assessment instrument design, including sample questions, is described in detail in Rivet & Kastens (2012).

Data collection and analysis

The assessments activities were administered in a pre/post format to all of the sections of each of our partner teachers in both Year 1 and Year 2 of the study, for a total of 357 consenting participants. As we were interested in understanding the relationship of assessment items as reflecting the proposed construct rather than student gain, we examined both the pretest and posttest together in the same data set. Therefore, the total number of

cases considered was 707, with 323 cases from year 1 and 384 cases from year 2. An expanded outcome space was used to map each item response made by the student to a particular level of the construct. A kappa calculation of .84 for moon, .86 for seasons, and .90 for deposition provided a strong sense of inter-rater reliability. Prior analysis (Rivet, et al., 2013) demonstrated that the three assessments were generally comparable at measuring student reasoning around the levels of the construct, providing validity for claims drawn by looking across assessments in the three different topic areas.

With parameters estimated using R software, the assessment data was analyzed using a Rasch modeling approach for polytomous data. This approach provides an estimate of student ability and test item difficulty, both of which can be approximated based on the overall performance of a given sample of students on an instrument (Wilson, 2005). A Cronbach's alpha of .81 was established for the multidimensional analysis across the three topics, and the expected *a posteriori* (EAP) was 0.84 for moon assessment items, 0.82 for seasons, and 0.80 for deposition. Ability estimates were calculated for each student on every test they completed to determine their proficiency score in relation to the difficulty of the items. The average ability estimate for moon posttests was 0.51, or about a level 2 non-correspondence, 0.25 for seasons (level 2 correspondence), and -0.35 for deposition (level 2 correspondence). For each posttest, the students' ability estimate was utilized to assign a level of analogical reasoning based on our construct. Levels were determined by taking the average ability estimate score from the two adjacent levels, which produced a threshold score. For example, the threshold score was calculated by averaging the mean score of level 1 correspondence (1c) and mean score of level 1 non-correspondence (1n) within a topic. This average score established the cut-off point for a student to receive either a level 1 correspondence or level 1 non-correspondence score.

Of the 357 students included in the above analyses, 29 were identified as target students for further examination. These target students had completed all six tests (except for one group of six students from year 1 that did not receive a seasons pretest) and had participated in at least two interviews from the three topics. The students' demographics are outlined in Table 1. In examining the ability estimates of the target students' posttests in comparison to the overall data set, these students appear to be representative of the larger student population included in this study.

Table 1: Target student demographics.

Target Student Demographics (n_{total} = 29 students, n_{year1} = 16 students, n_{year2} = 13 students)						
School	MS1 = 11	MS2 = 8	MS3 = 2	HS1 = 8		
Teacher	T1 = 8	T2 = 8	T3 = 5	T4 = 3	T5 = 2	T6 = 3
Level	Lower = 1		General = 20		Advanced = 8	
Gender	Female = 11		Male = 18			

The target student interviews were transcribed and analyzed along several dimensions, including the robustness of their articulated conceptual understanding of the causal mechanism driving each of the modeled Earth phenomena (moon phases, seasons, and deposition). Student responses in each interview were rated on a scale from weak to excellent. Weak understanding included descriptions of relative motions or configurations of entities in the Earth System, but no connections to causes or mechanisms of the emergent phenomena. A good understanding was rated when there was a single cause or preliminary mechanism given, whereas an excellent understanding indicated the influence of multiple coordinated factors on the resulting phenomena. The specific criteria for each level were tailored to the particular content focus of the interview. Each target student received a single rating score for each topic based on their responses across questions in the interview. We then compared target students' content scores from the interviews with their posttest analogical reasoning level ascertained from the written assessment for each topic. These ratings were categorized to identify clusters of students that shared similar profiles of content understanding and analogical mapping across content areas.

Findings

Overall, there were a few observed trends across the target student group focused on in this analysis. First, based on the analogical mapping proficiency scores calculated by the assessments, none of these students were found to reason only at the lowest measureable level, that of only being able to map correspondences between the model and the Earth System at Level 1: Entities and Attributes. Everyone in this group was also able to at minimum reason around both correspondences and non-correspondences at this level for every topic. The Moon assessment demonstrated the greatest range of proficiencies along the construct, with four students demonstrating proficiency at only Level 1 correspondences and non-correspondences, and eleven students demonstrating proficiency at mapping non-correspondences at Level 3: Causation and Mechanism, the most sophisticated reasoning measured on the written assessment. The other two assessments, focused on seasons and deposition, were more challenging for students. None of the students in this group were found to

consistently demonstrate proficiency at Level 3 for either of these two topics. That is not to say that they did not occasionally respond accurately to assessment items targeted to these levels. Rather, with respect to the whole student population who completed these assessments, the target students were not the most proficient at sophisticated Level 3 mapping in the seasons and deposition assessment activities.

We identified four clusters, or profiles, of the relationship between the sophistication of students' analogical reasoning around science models and their content understanding as articulated during the interviews. Each of these profiles is described in detail below.

Profile 1: "Level 1" Mapping with Weak Content

The first pattern that emerged were a relatively small group of target students (3 of the 29, or 10%) whose average proficiency rating on each of the three assessments indicated overall ability to articulate correspondences and non-correspondences only at the "Level 1: entity and attribute" stage of the construct map. These students also demonstrated limited understanding of the science concepts under study consistently across the interviews. The combination of relatively weak content understandings and limited proficiency at identifying and articulating the relationship between the model and the Earth System phenomena is to be expected as students are just beginning to develop fluency with both the science concepts and modeling practice. We would consider such students to be near the beginning of a progression that describes this kind of learning. What is also of note in these groups of students is the consistency of their typical responses across topics, both in terms of mapping between the model and the Earth System and the sophistication of their descriptions of the science phenomena itself. Unlike other profiles identified, described in more detail below, students' performance in both the three posttest assessment tasks and the two or three interviews they participated in was consistently near the lower end of both the content and the analogical mapping scale.

One student who fit this profile, Stephanie, demonstrated proficiency at mapping between the model and real Earth System only at a Level 1 during the written assessment, and no higher than a Level 2 during the interview. For example, during the Moon Phase interview she said that the model was like the phases of the moon because "as the moon revolves around, it should show the changes." Epistemologically, such comments reflected a common perception amongst students in this profile that the model "shows" the science concept that it is intended to represent, without indicating clearly what that science concept was or how the model represented that idea. Stephanie did not make any statements to suggest why we see phases beyond explaining that it relates to the revolution of the moon around earth. She showed similar limitations for the topic of seasons. Her poor content knowledge and limited mapping ability did not deter Stephanie from using the physical model to communicate, as she was observed using the physical model to illustrate both accurate and inaccurate conceptions about the entities in the real Earth System and the relative motions of these entities.

Profile 2: Medium Proficiency for Content Understanding and Analogical Mapping

A second profile that emerged from the target student analysis was that almost a quarter of these students, 24%, were generally in the middle in terms of both sophistication of mapping the correspondences and non-correspondences between the model and the real Earth System, and the robustness of their content understanding. This pattern also reflects the ability level achieved by the majority of students in the larger study. As students are moving through the Earth Science curriculum in middle and high school, there is evidence that they have a better understanding of both the ways that scientific models represent, and can be used to understand, specific Earth Systems phenomena than they did in the lower grades. Specifically, they generally characterize the phenomena under study in terms of spatial relationships between entities, and are able to accurately identify and describe Level 2: Configuration and Motion correspondences between the model and the phenomena. However, these students still struggle with characterizing the non-correspondences of motion or configuration in the model, and are limited in their explanations of the mechanism behind emergent phenomena across topics. We consider this an appropriate "stepping stone" understanding (Wiser, Smith & Doubler, 2012) as students develop increased fluency with both modeling and Earth System understandings.

One student who fit this profile, Robert, demonstrated consistent Level 2 mapping and good, but not excellent content knowledge. During the seasons interview, Robert attributed the changing seasons to the tilt of the Earth's axis in different positions in its revolution around the sun, but did not explain more sophisticated aspects such as the effect that the changing angle of light has on the amount of solar energy received at a point on Earth. Robert was able to explain his mapping between the real Earth's revolution and tilt and that of the model, but illustrated tilt inconsistently using the physical model. During the deposition interview, Robert explained how the attributes of sediment might influence settling rates but conflated density and size. Robert never used the physical model for deposition, in contrast to the seasons model which he used frequently.

Profile 3: Robust Content and Mapping Fluency

A rather surprising result of our analysis was the finding that over 40% of the target students (12 of the 29 interviewed) demonstrated robust content understanding of the science concepts across the three topics, and

relatively high sophistication in terms of their ability to map correspondences between the model and the real Earth System in the assessment activities. As described earlier, the assessments for the seasons and deposition topics were limited in the extent to which evidence of Level 3: Causation and Mechanism correspondences, and particularly non-correspondences, was measured. There was also observed a relative range of both mapping proficiency and robustness of content explanations in this group. However, when looking across the three topics on their posttest and post-instruction interview responses, it was evident that a majority of our target students were able to demonstrate both modeling and conceptual understanding with an appropriately expected level of sophistication given the population.

One student, Emily, demonstrated both strong knowledge of each topic and a strong ability to explain how the model illustrated the real system at various levels. Emily did not use the physical model provided during the interviews frequently; however, she did use it to communicate sophisticated ideas and demonstrate Level 3 mapping.

Profile 4: Consistency of Analogical Mapping Ability with Variable Content Understanding Across Topics

Seven students in our target group (24%) did not fit into any of the three prior categories. Four of these students (14% of the total group) demonstrated an interesting pattern in that the sophistication of their analogic mapping between the model and the Earth System demonstrated through the written assessment was consistent across topics, yet during the interview they displayed vastly different levels of conceptual understandings of the science phenomena under study. Three of these students were proficient at the Level 2 mapping, however demonstrated excellent understanding of either moon or seasons phenomena but poor understanding of deposition. One of these students showed a similar trend in terms of content, but mapped at the lowest level across the written assessments.

One student, Sarah, demonstrated poor mapping consistently despite variations in the robustness of her content knowledge across the topics. Sarah demonstrated Level 1 mapping along with poor knowledge of deposition. For Moon Phases, Sarah was able to explain verbally, and with the aid of a physical model, what the causes are for the phases of the moon. However, she was not able to map between the model and real Earth for more specific questions relating to particular configurations of the model and how these mapped to real phases. For questions such as these, Sarah struggled to explain her reasoning.

Outlier Examples

The three students who did not fit into any of the above categories showed unique patterns of mapping and content understanding across the three topics. One student consistently mapped at the mid to upper range of ability levels for a Level 2, but demonstrated poor content understanding across the board. A second student demonstrated consistently excellent conceptual understanding across topics, but performed at the lowest level of analogical mapping consistently across the assessments. The third student in the group showed a pair-wise trend: excellent content understanding of moon and high sophistication of mapping; a level 2 mapping ability on seasons (no interview); and a lower conceptual understanding of deposition with proficiency of analogical mapping on this assessment at a Level 1. To our surprise, this was the only student out of the 29 target students interviewed who showed this trend.

One example of a student in this group is Michael, who showed difficulty applying the knowledge he had of the causes of the seasons to a model, both on the written assessment and during the interview. This student never used the provided physical model of seasons during his explanations, and when describing how the model was like the seasons, he gave a response that was only at a Level 2, stating "the earth goes around the sun and it is rotating [like] the ball orbiting around the [basketball], the seasons change in different areas." However, throughout the interview Michael showed that he understood more sophisticated aspects of the causes of the seasons including the influence of the tilt of Earth on the angle of sunlight, which he explained affected the amount of sunlight an area received.

Discussion: Bootstrapping Towards Conceptual Understandings

The analysis of target student profiles illustrates some significant characteristics that demarcate the evolving reciprocal relationship between understanding Earth System models *as* scientific representations, and understanding the science concepts *of* the represented entities, configurations, motions, and emergent phenomena in the real Earth System. First, it is notable that none of the target students included in the analysis were capable of mapping at any significant level without robust content understanding as well. Our findings demonstrate that it is not the case that analogical mapping is a generalized and transferrable context-independent skill. If it were, it is possible that we would have observed students engaged in sophisticated analogical reasoning while also demonstrating weak conceptual understanding. Rather, our data support the claim that analogical reasoning only exists *in the context of the phenomena and concepts that are being reasoned about*. This indicates that the power of bootstrapping between representations and phenomena to develop robust

conceptual understandings must begin with at least some limited understanding of the phenomena under study to initiate the bootstrapping process.

A second important trend observed in this analysis is that the sophistication of analogical mapping around models and the robustness of students' conceptual understanding co-vary across the profiles. Overall 22 of the 29 target students (over 75%) were consistent in terms of falling into either a "low mapping/low content", "medium/medium", or "high/high" profile. This finding in particular supports the claim that increased sophistication in modeling practice and increased sophistication of conceptual scientific understanding do indeed co-vary, and exist together in a reciprocal developmental relationship. This lends support to the power of considering bootstrapping as a productive mechanism to describe the relationship between developing modeling and conceptual understandings.

This research points to both specific recommendations for instruction and curriculum design, and areas in need of further research. By recognizing the nature of the reciprocal relationship between content understanding and modeling practice, this work encourages curriculum and instruction to avoid teaching science concepts and modeling practices as separate knowledge domains. Additionally, the profiles themselves illuminate the nature of both students' engagement and use of models and the sophistication of their conceptual reasoning at each of these levels. Such information can be used by teachers to help assess and support student learning within and across content areas. Next steps in this work include further analysis of science classroom environments engaged with robust model use, with the aim of understanding the various ways that instruction can influence the relationship between students' modeling practices and concurrent development of science content understanding.

References

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, *18*(1), 32-42.

Carey, S. (2004). Bootstrapping and the origin of concepts. Daedalus, Winter, 59–68.

Gentner, D. (1983). Structure-Mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155-170.

Kastens, K., & Rivet, A. (2010). Using analogical mapping to assess the affordances of scale models used in earth and environmental science education. In C. Holscher et al. (Eds.), Spatial Cognition VII, LNAI 6222 (pp. 112–124). Berlin: Springer-Verlag.

Koslowski, B. (1996). Theory and evidence: The development of scientific reasoning. Cambridge, MA: MIT Press.

Kurtz, K. J., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, *10*(4), 417-446.

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, D.C.: The National Academies Press.

Nersessian, N. J., & Chandrasekharan, S. (2009). Hybrid analogies in conceptual innovation in science. *Cognitive Systems Research*, *10*(3), 178-188.

Nersessian, N. J., Kurz-Milche, E., Newsletter, W. C., & Davies, J. (2003). Research laboratories as evolving distributed cognitive systems. In R. Alterman & D. Kirsh (Eds.), Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society (pp. 857–862). Mahwah, NJ: Lawrence Erlbaum.

Rivet, A. E. & Kastens, K. (2012). Developing a construct-based assessment to examine students' analogical reasoning around physical models in earth science. *Journal of Research in Science Teaching, 49*(6), 713-743.

Rivet, A., Lyons, C., Miller, A., Schmalstig, M. & Kastens, K. (2013). Exploring students' reasoning around models in Earth Science. Paper at the annual meeting of the National Association for Research in Science Teaching. Rio Grande, Puerto Rico: April 2013.

Schwarz, C., Reiser, B., Davis, B., Kenyon, L., Acher, A., Fortus, D., et al. (2009). Designing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching, 46*(6), 632-368.

Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, *23*(2), 165-205.

Wilson, M. (2005). Constructing measures: An item response modeling approach. Mahwah, NJ: Erlbaum.

Wiser, M., Smith, C. L., & Doubler, S. (2012). Learning Progressions as Tools For Curriculum Development. In *Learning Progressions in Science* (pp. 359-403). SensePublishers.

Acknowledgments

We would like to thank Kim Kastens and Mariana Schmalstig for their assistance with this research. The research reported here was supported in part by the National Science Foundation # DRL09-09863 and DRL09-09982. All opinions expressed are those of the authors and do not necessarily represent either the funding agency or Teachers College Columbia University.

Towards a Complex Systems Meta-Theory of Learning as an Emergent Phenomenon: Beyond the Cognitive Versus Situative Debate

Michael J. Jacobson, The University of Sydney, Australia, michael.jacobson@sydney.edu.au
Manu Kapur, Nanyang Technological University, Singapore, manu.kapur@nie.edu.sg
Peter Reimann, The University of Sydney, Australia, peter.reimann@sydney.edu.au

Abstract. This paper proposes a meta-theory of learning based on conceptual perspectives and methodologies being employed in the study of complex physical and social systems to inform research in the learning sciences and education. The contexts in which learning occurs are in fact complex systems with elements or agents at different levels—from neuronal, cognitive, intrapersonal, interpersonal, cultural—in which there are feedback interactions within and across levels of the systems so that collective properties arise (i.e., emerge) from the behaviors of the parts, often with properties that are not exhibited by those parts. We analyze the long running cognitive versus situative learning debate and propose that a *complex systems meta-theory of learning* (CSMTL) provides a principled way to achieve a theoretical rapprochement. We close by considering other theoretical and methodological implications of the CSMTL for research in the learning sciences.

There are various perspectives from which to ground systematic inquiry into *learning*, which, of course, is the central enterprise of educational research. Discussions of these perspectives tend to argue for the primacy of a specific locus of theory and philosophy that in turn grounds various research agendas that generally intend to validate, enhance, or challenge particular perspectives. As a field that studies learning and education, there have important debates or "fault lines" (diSessa, 2006) about theory and methods. For example, there has been a vigorous debate of theoretical import about the primacy of cognitive (i.e., individual) versus situative (i.e., socio-cultural) perspectives about learning (Anderson, Reder, & Simon, 1996, 1997; Greeno, 1997; Norman, 1993). We are perhaps on the verge of a third distinct theoretical and empirical perspective about learning that is emerging the neurosciences, and there are already appeals to assigning primacy for theory and research about learning from this field over cognitive and socio-cultural perspectives (for a critical discussion, see Bruer (2006)). As another example, there have been the so called "methodology wars" between proponents of quantitative and qualitative methods, with mixed methods employing elements of both representing perhaps an uneasy truce between these camps (e.g., Firestone, 1987; House, 1991).

Unfortunately, these debates related to theory and methodologies in the study of learning and educational research have been persisting for decades. This inability of the field to reconcile or vindicate one camp or another is a serious issue. For researchers, this has meant a "community of practice" in educational research that has fractured into "cognitive," "socio-cultural," and neuro-science" silos that are theoretically and methodologically isolated from each other, or perhaps worse, that simply ignore each other.

Given such debates seem to involve a "clash of cultures" (Norman, 1993, p. 3), how might a rapprochement be made? We argue in this paper that conceptual perspectives and methodologies being employed in the study of complex physical and social systems may help reconcile certain existing debates in the field and to help provide an enhanced foundation to use in educational research more generally. The paper is organized into four main sections. First, we provide a brief overview of the cognitive versus situative theories debate. In the second section, we propose an initial set of components for a *complex systems meta-theory of learning* (CSMTL). In the third section we discuss ways in which the CSMTL provides a principled reconceptualization and rapprochement of the cognitive-situative debate and related issues in the field. We close the paper with suggestions for future research involving the CSMTL and implications for the field of educational research more generally.

Cognitive Versus Situative Theories: An Overview of the Debate

In the seminal paper advocating a situated perspective of learning, Brown, Collins, and Duguid (1989) argued that knowledge should be viewed as situated, as being a "product of the activity, context, and culture in which it is developed and used" (p. 32). Such a perspective had important implications for schooling, which they believe had been narrowly concerned with the transfer of abstract and decontextualized formal concepts. However, the cognitive science research upon which many of the key arguments for situated learning by Brown, Collins, and Duguid was itself generating considerable debate. In 1993, a special issue of *Cognitive Science* pulled together nine papers that debated two perspectives about the study of human cognition. In the introductory paper by Norman (1993), he framed these two perspectives—the traditional symbolic approach for studying human cognition that focused on the processing structures and symbolic representations of the brain and the "new

 © ISLS

upstart, the study of situated cognition" (also referred to as "situated action" or "situativity" by other authors) that focuses on the structures of the world constraining and shaping human behavior. The primacy of a cognitive level of analysis was clearly championed by Vera and Simon (1993), who argued that symbolic models of individual cognitive processes and representations have been quite successful in providing principled accounts of humans and their interactions with the world. In contrast, in arguing for the primacy of the situative level of analysis, Greeno and Moore (1993) propose the term "situative" to describe cognitive processes as interactions between a person and other people and physical systems. Similarly, Suchman's (1993) notion of situated action emphasizes "constructing accounts of relations among people, and between people and the historically and culturally constituted worlds that they inhabit together" (p. 71).

This debate broadens and deepens in many ways as reflected in a series of papers in *Educational Researcher* in the middle to late 1990s in which the focus shifts from considerations of how people think and act to implications of cognitive versus situative perspectives for teaching and learning. The paper of Anderson, Reder, and Simon (1996) characterized situated learning as a view that much of what students learn is specific ("situated") to the context in which it was learned, which implies knowledge does not transfer between tasks and that learning abstractions is of little value. They go on to provide a critique of the application of situated learning in mathematics education in particular, and propose that educational approaches based on cognitive research into learning processes may be more efficacious than those based on situated perspectives.

The following year, Greeno (1997) provided a response in which he argued that the main differences between situative and cognitive perspectives discussed by Anderson, Reder, and Simon were primarily due to underlying framing assumptions of these two perspectives. In particular, Greeno (1997) maintained:

> The cognitive perspective takes the theory of *individual cognition* [italics added] as its basis and builds toward a broader theory by incrementally developing analyses of additional components that are considered as contexts. The situative perspective takes the theory of *social and ecological interaction* [italics added] as its basis and builds towards a more comprehensive theory by developing increasingly detailed analyses of information structures in the contents of people's interactions. (p. 5)

In the same issue of *Educational Researcher*, Anderson, Reder, and Simon (1997) provide a rejoinder to Greeno in which they found a degree of agreement between the cognitive and situative positions on evidence for findings and a consensus on certain educational issues. They also agreed that Greeno raised a substantive issue as to "whether the more profitable research path is one that takes individual or social activity as the principal unit of theoretical focus" (p. 20). Not surprisingly, Anderson et al. end their rejoinder with a robust assertion of the superiority of the cognitive information processing approach over a situative theoretical approach.

This debate broadens in a paper by Cobb and Bowers (1999) in which they criticized the conflicts between cognitive and situative learning theories as being of primary interest to educational psychologists and not to educators involved with classroom-based learning design and research. Still, the detailed discussion of their research for studying the learning of mathematics in classrooms primarily employed a situative analysis approach as they found little theoretical utility in the cognitive perspective of Anderson, Reder, and Simon for understanding the "essence of individual and collective human activity" (Cobb & Bowers, 1999, p. 13).

In 2003, Derry and Steinkuehler provided a critical review of the literature related to cognitive and situative theories. They proposed that cognitive theory regards cognition as symbolic computation, and broadly includes perspectives of *socio-cognitive* theoreticians such as Piaget as well as others summarized by Anderson, Reder, and Simon (1996). The situative perspective according to Derry and Steinkuehler embraces a family of social science theories including *situated cognition*, *sociocultural theory*, *distributed cognition*, and *activity theory*. Derry and Steinkuehler propose what might be called a "pragmatist view" of the cognitive-situative debate, as they comment that many researchers and designers working in classroom environments were fusing points of view from the cognitive and situative perspectives. However, they also note that a well-defined theory between these two communities of educational practice had not been proposed, which we believe is still true today. They speculate that what may emerge is a:

> *complex systems theory* [italics added] of cognition understood in its broadest ecological sense, and that the resulting methodological approach will be superior to either theoretical viewpoint standing alone, capable of providing more complete understanding of learning and education. (Derry & Steinkuehler, 2003, p. 805)

We next describe our initial efforts in this area, which we regard as a meta-theory of learning based on complexity.

Towards a Complex Systems Meta-theory of Learning

In this section we outline a complex systems meta-theory of learning that may provide a principled basis for a rapprochement between cognitive and situative perspectives, as well as to inform other issues in the field. However, what do we mean by a *meta-theory*? As the name suggests, a meta-theory is a theory about theories. While a theory is concerned with specifying concepts and relations that can provide accounts for aspects of the natural or social world, a meta-theory provides concepts for describing what form theories should take, and for identifying requirements of what they should achieve. We also note that at this time there is not a general "theory" of complex systems. Rather, the study of complex physical, biological, and social systems by multidisciplinary fields has been providing a framework of conceptual perspectives, principles, and methods (e.g., emergence, sensitivity to initial conditions, dynamical attractors, agent-based modeling, scale-free networks) that we believe can function to generate and evaluate specific theories of relevance to particular types of systems (Jacobson & Wilensky, 2006).

We conceive of learning not as something that *is*, but rather, as something that *emerges* in the context of formal and informal systems of education. We concur with Clancey (2008) that environments in which human cognitive processes—and therefore learning—occur are in fact *complex systems* that are "inherently social, interactive, personal, biological, and neurological, which is to say that a variety of systems develop and depend on one another in complex ways" (p. 11). As we discuss in more detail below, the construct of *emergence* is a central one in the study of complex systems, which encompasses several new theoretical perspectives (e.g., general systems theory, complexity theory, system dynamics, complex adaptive systems, chaos theory) as well as attendant approaches for modeling these systems (Bar-Yam, 2003; Gell-Mann, 1994; Gleick, 1987; Holland, 1995; Mitchell, 2009; Prigogine & Stengers, 1984) that represent important methodological innovations (also of relevance to learning sciences research). As background, we next provide an overview of perspectives about complex systems and complexity and then rejoin our consideration of the main components of a complex systems meta-theory of learning.

What is Complexity?

Scientific study of complex systems—sometimes referred to as *complexity*—over the past three decades has lead to insights about the world that classical approaches tended to over simplify or to ignore (Bar-Yam, 2003). Briefly, complex systems consist of elements or agents that interact with each other and their environment often based on simple rules. Feedback interactions within and across levels of the system result in self-organization, with emergent patterns forming at mezzo and macro levels of the system. There is also a dialectical co-existence of linearity and nonlinearity in the behavior of complex systems, such as the linear predictability of seasons that emerges out of the nonlinear and probabilistic nature of day-to-day weather. Another key characteristic of complex systems is that collective properties arise (i.e., emerge) from the behaviors of the parts, often with properties that are not exhibited by those parts. Examples of complex systems include adaptation of white blood cells to invading bacteria, emotional and cognitive brain behaviors out of the interaction of individual neurons, the flocking formation of individual birds, dynamic equilibrium in ecosystems out of individual predator-prey interactions, segregation patterns in cities out of individual choices in places to live, and so on.

However, before we can advance our argument for a complex systems meta-theory of learning, the question may be asked if the study of complex systems has yielded findings or insights that are different than those from theoretical, research, and disciplinary perspectives of more traditional scientific fields such as physics, biology, chemistry, and so on. This is an issue discussed in the recent book by Mitchel (2009). One significant contribution is the conceptualization of complex problems in ways that challenge long-term scientific assumptions. As examples, chaos has demonstrated that intrinsic randomness of a system may not be necessary for the overall behaviors of the system to look random; recent findings in genetics challenge the centrality of the role of genetic change in evolution; and chance and self-organization are being viewed as dynamics that challenge the primacy of natural selection in evolution. Mitchel also notes the importance in both scientific communities and the general population of ways of thinking that include nonlinearity, decentralized control, networks, hierarchical levels in systems, statistical representations of information, and so on. We next consider how selected complexity ideas such as these are now being incorporated into educational research.

Research on Learning and Complex Systems

There has been a shift in the learning sciences and related fields of educational research over the past decade from earlier work on students learning concepts about complex systems to the application of perspectives about complex physical and social systems to understanding learning processes and environments (for an overview, see Jacobson and Wilensky (2006)). One indication of this latter trend is reflected in the use of complexity concepts by researchers who are studying learning environments. For example, Bereiter and Scardamalia (2005) have argued that:

> As complex systems concepts such as self-organization and emergence make their way into

> mainstream educational psychology, it becomes increasingly apparent that there are no simple causal explanations for anything in this field. In general, what comes out of a sociocognitive process cannot be explained or fully predicted by what goes into it. Creative works, understanding, and cognitive development are all examples of complex structures *emerging* [italics added] from the interaction of simpler components (Sawyer, 1999, 2004). Learning itself, at both neural and knowledge levels, has emergent properties (p. 707) (Pribram & King, 1996).

The critique of simple causal explanations made by Bereiter and Scardamalia centers on the construct of *emergence*, that is, properties emerging from the interaction of simpler components. We believe that the construct of emergence is centrally important for the study of learning that has important theoretical and methodological implications. Before considering these implications, we next "unpack" emergence as well as related perspectives of linearity and nonlinearity, more fully.

Emergence and the Dialectics of Linearity and Nonlinearity (1)

Interest in emergence is a recent area for learning and cognitive scientists (Clancey, 2008; Goldstone, 2006; McClelland, 2010). For example, inter-subjective processes at the local (individual) level yield cognitions—such as opinions (Isenberg, 1986), generation of abstract representations (Schwartz, 1995), representation and schema learning (Rumelhart, Smolensky, McClelland, & Hinton, 1986), group dynamics (Kapur, Voiklis, & Kinzer, 2008), knowledge building (Bereiter & Scardamalia, 2005), among others—that differ both in complexity and kind from those produced by any collaborating agent or those expected from the central tendency among collaborators (Vallabha & McClelland, 2007). Moreover, these cognitions emerge spontaneously, without forethought or awareness among collaborating agents (Goldstone, 2006). Apparently, both the individual and the group learn. Complexity theory posits that learning is at once distinct and emergent—which is consistent with the critique of simple causal mechanisms by Bereiter and Scardamalia above.

However, the concept of emergent behavior is paradoxical. On the one hand, it arises from the interactions between agents in a system (e.g., individuals in a collective). On the other hand, once such a behavior emerges in a system, it influences and/or constrains subsequent interactions between the micro level actions of agents. Thus an emergent pattern or behavior can seem to have a life of its own independent of the local interactions (Kauffman, 1995) and therefore, cannot be reduced to the individual agents (or parts) of the system (Lemke, 2000). For example, a traffic jam emerges from the local interactions between individual drivers; at the same time, it constrains the subsequent local interactions between these individuals. Once underway, traffic jams do seem to have a life of their own—such as the backwards propagation of a traffic jam (i.e., a clump of cars)—and this emergent pattern cannot be reduced to the behavior of the individual cars that generally move forward. Similarly, structures (norms, values, beliefs, lexicons, and so on) within social networks emerge from the local interactions between individual actors, and then, once emerged, these structures constrain the subsequent local interactions between these actors (Lemke, 2000; Watts & Strogatz, 1998).

Understanding *how* properties such as opinions, representations, group dynamics, indeed, learning actually emerge requires, in our view, a careful consideration of how macro-level behaviors emerge from and constrain micro-level interactions. This critical dynamic relates to another important complexity perspective—*the co-existence of linearity and nonlinearity in complex systems,* which we illustrate with an example. Consider the brain as a collection of neurons (agents). These neurons are complex chemical systems themselves, but they exhibit simple binary behavior in their synaptic interactions that are often modeled as *linear*, probabilistic functions (or rules). This type of emergent behavior, which demonstrates the complexity at the individual micro level resulting in simplicity at the collective mezzo level, is called *emergent simplicity* (Bar-Yam, 2003). Put another way, *nonlinear* chemical reactions can result in a *linear*, global behavior.

Linearity may be broadly conceived both as a *mathematical operator* as well as a *functional relationship*. A linear operator is essentially an additive operator (Bertuglia & Vaio, 2005). For example, traditional analytical methodologies such as linear differential equations and statistical modeling, regardless of their mathematical sophistication, are essentially linear operators. They work well for closed, linear systems (or approximations thereof) where the whole is equal to the sum of its parts, thus allowing one to break a system into its components or parts, study the parts individually, and then add the parts together to form the whole. However, applying the linear operator and its associated quantitative methodologies to the study of emergent behavior in open systems is *fundamentally problematic* because, by definition, *emergent properties have properties that are different than the parts, not additively composed of the parts*.

Linearity may also be conceived as a functional relationship, such as constant proportionality or a straight line. When applied to model a causal relationship, linearity restricts one to phenomena in which the effects are proportional to their causes. This is because linearity tends to treat small changes or perturbations as temporally transient without any long-term effects. However, emergent behavior often exhibits *nonlinear* global

effects *even if the local action is linear*. As we demonstrated above, linearity and nonlinearity can co-exist in a system, and thus one cannot assume that global effects are proportional to their local causes. In fact, small linear changes or perturbations such as El Nino can and often do have large, nonlinear effects. Nonlinearity in climate systems has been metaphorically referred to as the so called "butterfly effect" in the seminal research of Lorenz (1963) a half a century ago. Unfortunately, important nonlinear relationships among variables across scales and hierarchies may be missed entirely, or worse, be inappropriately and inaccurately modeled linearly since that is only what the linear method can handle (Holland, 1995).

However, it is *not* the case that understanding emergent behavior requires that we make a "conceptual shift" from linearity to nonlinearity. Rather than characterizing the distinction between simple and complex systems in terms of dichotomies such as linear versus nonlinear, we propose a more productive characterization of complexity, and of emergent behavior, lies not in emphasizing these dichotomies, *but in collapsing them*. We argue that complexity is better characterized as a *dialectical co-existence of linearity and nonlinearity*. The complexity of emergent behavior comes from the co-existence of linearity and nonlinearity across and within multiple levels or scales of an open system. Indeed, because of this, complex systems exhibit seemingly opposing properties and behaviors: randomness and order, predictability (e.g., attractors, highly connected nodes or hubs) and unpredictability, coherence and incoherence, stability and instability, centralization and decentralization, and so on. Complexity is not one or the other, it is *both* (Kauffman, 1995).

There are important general implications of emergence, linearity, and nonlinearity for understanding the dynamics of learning and cognition in complex social systems. Perhaps the most important implication has been well summarized by Gureckis and Goldstone (2006): "Rules that govern behavior at one level of analysis (the individual) can cause qualitatively different behavior at higher levels (the group)" (p. 1). We revisit this critical point in our discussions of debates such as cognitive versus situative perspectives and quantitative versus qualitative methodologies for educational research.

The CSMTL and Re-conceptualizing the Cognitive Versus Situative Debate

Recall that the central argument of this paper is that there are key long standing educational research debates that we believe may be re-conceptualized through the use of a complex systems meta-theory of learning (CSMTL) that is based conceptual perspectives and methodologies employed in the study of complex physical, biological, and social systems. To illustrate this thesis, we focus on the debate about the theoretical primacy of cognitive versus situative perspectives on human cognition and learning. In our review of the cognitive-situative debate, there are two main aspects of contention in the literature: (a) *level of theoretical primacy and theoretical mechanisms*, and (b) *methodologies for research*. We also believe there is third issue implicit in this debate: (c) *epistemic challenge of simple explanations for complexities of learning*. Given space limitations, we next consider (a) and (c) here; please see Jacobson and Kapur (2012) for a discussion of (b).

Level of Theoretical Primacy and Theoretical Mechanisms

In the cognitive-situative debate in the 1990s, there was a clear advocacy of theoretical primacy for one perspective or the other that then relegated the other perspective to secondary importance. For cognitive advocates, individual cognition was the fundamental level and the social context was viewed as a secondary additional component (Vera & Simon, 1993), whereas situative advocates regarded the level of social and ecological interactions as being theoretically primary and individual cognition as secondary (Greeno, 1997).

A metaphor for this facet of the debate is that advocates of cognitive and symbolic representations view a system of learning as "trees" whereas advocates of situative perspectives view such a system as a "forest." Regarding a learning system in terms of its micro-level "trees"—i.e., individuals as cognitive symbolic processing agents or neurons as agents—provides essential details but will miss patterns that arise when trees are contextualized at a macro-level "forest"—i.e., classroom community or socio-cultural milieu. Conversely, focus on "forests" of situated social and cultural contexts will fail to see the details of the micro-levels of a learning system. Thus for this critical facet of cognitive-situative debate it follows, in our view, that there is not a sufficient argument for the theoretical primacy of *either* perspective.

But how might the CSMTL resolve this issue? As Bar-Yam (2002) has pointed out, scientists who study complex systems do not conceptualize such systems as either "trees" or "forest" but rather as "trees-forest." In terms of the CSMTL, this dynamic relationship between micro- and macro-levels of a learning system is the study of learning as an emergent phenomenon. The *agents* in a learning system, whether neurons or symbolic cognitive constructs, interact with each other and their environments based on *rules* and typically *self-organize* through *within-level feedback* processes. These micro-level agent interactions yield *emergent patterns* such as thoughts and feelings or social norms and practices at hierarchically higher levels of the learning system. Critically, once higher system level patterns emerge, these may influence, shape, or constrain agent behaviors at lower levels through *across-level feedback* processes. For example, a student who cognitively makes sense (i.e., constructs a mental model with components similar to a scientific expert) of a predator-prey ecosystem and a chemical reaction as being at equilibrium will likely be better able to engage in science inquiry

activities with other students and in turn come to gain greater insights into other scientific knowledge and to come to enjoy engaging in other authentic types of scientific practices. The "trees-forest" metaphor also applies to the dialectical co-existence of linearity and non-linearity as one leads to the other and vice versa, just as micro-level actions lead to macro-level actions that in turn may constrain micro-level dynamics.

We argue that the CSMTL provides a way to resolve which theoretical level of a system to employ—neither cognitive nor situative—but rather *both*. Related, the CSMTL provides *theoretically principled mechanisms* (e.g., self organization, feedback, emergence) from which to understand how patterns in learning that situative perspectives are interested in *emerge* from the micro-level cognitive processes, as well as how situated and cultural contexts can influence the cognitive and symbolic aspects of learning. Although our discussion here is of necessity brief, we believe the CSMTL is superior to either theoretical perspectives alone and thus should be explored further in future learning research as well as to analyze existing learning research from its meta-theoretical perspective.

Epistemic Challenge: Simple Explanations for Complexities of Learning?

We have thus far considered conceptual implications of CSMTL for educational research, primarily using the cognitive versus situative debate as a crucible for these considerations. In this section we consider another issue, which is the epistemic implications of complex systems meta-theory and methods. A key, and perhaps counterintuitive, epistemic aspect of complex systems views is that the apparent complexity in the behavior of many complex systems may be described in terms of the interaction of system elements based on relatively simple rules (see also our discussion of *emergent complexity* and *emergent simplicity* above). This issue seems implicit in views of Simon (1996): "The central task of a natural science is to make the wonderful commonplace: to show that complexity, correctly viewed, is only a mask for simplicity; to find pattern hidden in apparent chaos (p. 1)." We call this the *simplicity-complexity* epistemic view.

The CSMTL represents an epistemic challenge to what we believe is a reasonably common epistemic view of *complexity-complexity*, which is that complex systems such as the ones educational researchers study must have "complex" explanations whereas simple systems would, of course, have simple explanations. Indeed, a *complexity-complexity* epistemic bias—and its corollary, a *simplicity-simplicity* epistemic bias—would seem to be obvious characteristics of "common sense." For example, a simple machine such as a pulley may be explained as a rope wrapped around a wheel to raise or lower something, whereas the behavior and operation of a complex machine such as a modern jet airliner could only be explained with complex concepts from physics (i.e., Bernoulli effect), engineering and materials science, business models to finance and maintain, and so on.

In our reading of the cognitive-situative debate literature, there have been two main ways in which what we characterize as an epistemic commitment to complex theories for human learning may be found reflected in perspectives of both cognitive and situative educational researchers (see especially Anderson et al. (1996), Cobb (1999), and Greeno (1997)). First, whether viewed from cognitive or social perspectives, human action and learning are complex, and second, sophisticated (i.e., complex) theory—whether cognitive or situative—is required to explain the complexity of human actions. We regard these views as being influenced by the *complexity-complexity* epistemic bias.

It is an important epistemic challenge of the CSMTL that we do not necessarily have to seek complex explanations for complex behavior; such behavior may very well be explained from the "bottom up" via simple, minimal information, such as utility function, decision rule, or heuristics contained in local interactions (Nowak, 2004). Of course, we recognize that future learning theory development may or may not align with a simplicity-complexity epistemic view, indeed, the authors of this paper debate this point amongst ourselves. Still, being aware of epistemic assumptions such as these has value to learning researchers, rather than an implicit acceptance of a position that might bias theory development or interpretation of data.

Still, we stress that the CSMTL does not hold that the rules of agents in complex systems are deterministic, depriving humans of any form of agency or deliberate, goal-directed activity. In fact, it is quite the opposite; these rules are context-sensitive, probabilistic, and sensitive to initial conditions (chaotic), and should be seen as explanatory constructs and relations developed by researchers to explain complex phenomenon. Once cognitive structures emerge through across level feedback mechanisms, these structures constrain the very linear, synaptic interactions between neurons that they emerged from (Epstein & Axtell, 1996; Kauffman, 1995). Further, a host of other co-evolving factors—social, cultural, and environmental—are also critical for behaviors such as cognition to emerge. Indeed, McClelland (2010, p. 753) argues:

> I don't think that anyone who emphasizes the importance of emergent processes would deny that planful, explicitly goal-directed thought plays a role in the greatest human intellectual achievements. However, such modes of thought themselves might be viewed as emergent consequences of a lifetime of thought-structuring practice supported by culture and education.

Conclusion

Before concluding our advocacy for a complex systems meta-theory of learning, we reflect on the oft referred to story of an individual who stops to ask a drunk at night prowling around on his hands and knees underneath a street light "What are you doing?" The drunk replies: "I'm hunting for my glasses." "But sir, they are not here; where did you lose them?" the stranger asks. "Over in the dark alley," says the drunk," but I can only see here."

In educational and learning sciences research, our "street lights" are our theories and methodologies, so that the cognitive versus situative debate might be metaphorically regarded as two different streetlights. We argue in this paper that viewing learning as emergence locates this phenomenon, at least partly, in the dark alley, hence our interests in new complexity-grounded theoretical constructs and methodologies that are being used to study complex physical and social systems (Jacobson & Wilensky, 2006; Goldstone, 2006). It is to be expected, of course, that new theoretical and methodological perspectives will invariably tend to generate more questions than answers. We both encourage and welcome this process, with hope that perspectives we suggest from the CSMTL might answer at least some claims for right questions

In the history of the physical sciences, new theories, such as Einstein's general theory of relativity that accounted for the precession of the perihelion of the orbit of Mercury, helped direct empirical research in physics to make an important new discovery that was inconsistent with earlier Newtonian theory. Likewise, new instrumentation and enabled research methods, such as the telescope for Galileo or particle accelerators for modern high-energy physics, invariably led to new theoretical breakthroughs. We hope our nascent CSMTL might provide conceptual perspectives that re-conceptualize issues such as the long-standing cognitive-situative debate in educational research, and that methodologies such as computational modeling techniques (e.g., agent-based modeling) might provide new instrumentation for researchers in the learning sciences.

Overall, we hope that principled considerations of learning as an emergent phenomenon in complex neural, cognitive, situative, social, and cultural systems will yield critically important insights of central relevance to our field that might not otherwise be possible with current perspectives and approaches. In addition, we believe viewing the environments in which learning occurs as complex systems provides researchers with powerful conceptual and methodological tools that are also being used by scientists in other areas of research. That there may be synergies of theory and methods between researchers in our field with scientists in other fields has the potential to enable more cross-disciplinary research as well as opportunities to more directly link findings from other fields to issues being explored by educational researchers and vise versa. We conclude humble and mindful of Einstein's famous admonition—"everything must be made as simple as possible, but not simpler"—as we articulate these first steps of a complex systems meta-theory of learning.

Endnotes

(1) For a fuller discussion of these and related issues, please see Jacobson and Kapur (2012).

References

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher, 25*(4), 5-11.

Anderson, J. R., Reder, L. M., & Simon, H. A. (1997). Situative versus cognitive perspectives: Form versus substance. *Educational Researcher, 26*(1), 18-21.

Bar-Yam, Y. (2002). General features of complex systems *Encyclopedia of Life Support Systems*. Oxford, UK: EOLSS UNESCO Publishers.

Bar-Yam, Y. (2003). *Dynamics of complex systems*. New York: Perseus Publishing.

Bereiter, C., & Scardamalia, M. (2005). Technology and literacies: From print literacy to dialogic literacy. In N. Bascia, A. Cumming, A. Datnow, K. Leithwood & D. Livingstone (Eds.), *International handbook of educational policy* (pp. 749-761). Dordrecht, The Netherlands: Springer.

Bertuglia, C. S., & Vaio, F. (2005). *Nonlinearity, chaos, and complexity: The dynamics of natural and social systems*. Oxford, UK: Oxford University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Bruer, J. T. (2006). On the implications of neuroscience research for science teaching and learning: Are there any? A skeptical theme and variations: The primacy of psychology in the science of learning. *CBE—Life Sciences Education, 5*, 104–110. doi: 10.1187/cbe.06–03–0153

Clancey, W. J. (2008). Scientific antecedents of situated cognition. In P. Robbins & M. Aydede (Eds.), *Cambridge Handbook of Situated Cognition* (pp. 11-34). Cambridge, MA: Cambridge University Press.

Cobb, P., & Bowers, J. (1999). Cognitive and situated learning prespectives in theory and practice. *Educational Researcher, 28*(2), 4-15.

Derry, S. J., & Steinkuehler, C. A. (2003). Cognitive and situative theories of learning and instruction. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (pp. 800-805). England: Nature Publishing Group.

diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 265-281). Cambridge, UK: Cambridge University Press.
Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, D.C.: Brookings Institution Press/MIT Press.
Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*(7), 16-21.
Gell-Mann, M. (1994). *The quark and the jaguar: Adventures in the simple and the complex*. New York: Freeman and Company.
Gleick, J. (1987). *Chaos: Making a new science*. New York: Viking Penguin.
Goldstone, R. L. (2006). The complex systems see-change in education. *The Journal of the Learning Sciences, 15*(1), 35-43.
Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher, 26*(1), 5-17.
Greeno, J. G., & Moore, J. L. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science, 17*, 49-59.
Gureckis, T. M., & Goldstone, R. L. (2006). Thinking in groups. *Pragmatics and Cognition, 14*(2), 293-311.
Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. Reading, MA: Addison-Wesley.
House, E. R. (1991). Realism in research. *Educational Researcher, 28*(6), 2-9, 25.
Isenberg, D. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology, 50*, 1141-1151.
Jacobson, M. J., & Kapur, M. (2012). Learning environments as emergent phenomena: Theoretical and methodological implications of complexity. In D. Jonassen & S. Land (Eds.), *Theoretical Foundations of Learning Environments* (Second ed., pp. 303-334). New York: Routledge.
Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the Learning Sciences, 15*(1), 11-34.
Kapur, M., Voiklis, J., & Kinzer, C. K. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers and Education, 51*, 54–66.
Kauffman, S. (1995). *At home in the universe: The search for laws of self-organization and complexity*. New York: Oxford University Press.
Lemke, J. J. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity, 7*(4), 273-290.
Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Science, 20*, 130-141.
McClelland, J. L. (2010). Emergence in cognitive science. *Topics in Cognitive Science, 2*(4), 751-770.
Mitchell, M. (2009). *Complexity: A guided tour*. New York: Oxford University Press.
Norman, D. A. (1993). Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science, 17*(1), 1-6.
Nowak, A. (2004). Dynamical minimalism: Why less is more in psychology. *Personality and Social Psychology Review, 8*(2), 183-192.
Pribram, K., & King, J. (Eds.). (1996). *Learning as self organization*. Mahwah, NJ: Lawrence Erlbaum Associate.
Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature*. New York: Bantam Books.
Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Parallel distributed processing models of schemata and sequential thought processes. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. II, pp. 7–57). Cambridge, MA: MIT Press.
Sawyer, R. K. (1999). The emergence of creativity. *Philosophical Psychology, 12*(4), 447–469.
Sawyer, R. K. (2004). The mechanisms of emergence. *Philosophy of the Social Sciences, 34*, 260–282.
Schwartz, D. L. (1995). The emergence of abstract dyad representations in dyad problem solving. *The Journal of the Learning Sciences, 4*(3), 321-354.
Simon, H. A. (1996). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.
Suchman, L. (1993). Response to Vera and Simon's situated action: A symbolic interpretation. *Cognitive Science, 17*, 71-75.
Vallabha, G. K., & McClelland, J. L. (2007). Success and failure of new speech category learning in adulthood: Consequences of learned Hebbian attractors in topographic maps. *Cognitive, Affective and Behavioral Neuroscience, 7*, 53–73.
Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science, 17*, 7-48.
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*(393), 440–442.

Adventures in Argument: Training in Argumentation Influences Student Resource Use in Collaborative Meaning Making

Julia Gressick, Indiana University South Bend, 1700 Mishawaka Avenue,
South Bend, IN 46634, jgressic@iusb.edu
Sharon J. Derry, University of North Carolina-Chapel Hill, CB# 3500, Peabody Hall,
Chapel Hill, NC 27599-3500, derry@unc.edu

Abstract: Argumentation is the primary pedagogical strategy employed in the online undergraduate course Human Abilities and Learning Online (HAL Online). We conducted a controlled *in vivo* experiment in this course to examine the effects, on collaborative meaning making, of providing direct training in argumentation early in the course. The performance of a group receiving the treatment, *Trained Argumentation with Modest Scaffolding* (TAMS), was compared with an ecological control group that did not receive argument training: *Emergent Argumentation with Modest Scaffolding* (EAMS). We hypothesized that argument training would influence how students attended to, used, and shared instructional resources as evidence to support explanations in collaborative meaning making. Results indicated that TAMS exerted strong influence on how deeply and thoroughly students processed, were accountable to, and integrated instructional resources.

Introduction

Argumentation as pedagogical practice is widely advocated for its potential to improve learners' conceptual knowledge and ability to reason in the domains of science, mathematics and social science (e.g., Cavagnetto, 2010; Kuhn, 2010; Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012). Although there is consensus concerning the value of instructional argumentation, the literature on this topic reveals a complex landscape of perspectives on how to conceptualize and design such instruction. In his review of argument in science education, Cavagnetto (2010) identifies three forms of argument as pedagogy: direct instruction in argument structure prior to engaging in scientific activity; developing argument skills through mentoring during immersion in science process; and instructing through ethical or political dilemmas, engaging argument processes that individuals experience socially from a young age (Hay & Ross, 1982). In agreement with Sandoval (2005) and others, Cavagnetto makes a case for the immersive approach. Like Kuhn (2010), however, he recognizes that more research is needed to understand this issue. Like others (e.g., Larson, Britt, & Kurby, 2009; Schworm & Renkl, 2007), we will make a case for direct, explicit instruction in argument, at least for contexts similar to ours.

The setting for the work reported here is Human Abilities and Learning, an upper-level undergraduate course offered at a large university. It is required for many majors, including teacher education. The course addresses the scientific basis of thinking and learning and what this implies for guiding children and adults, for personal development, and for building environments that help people learn and grow successfully. Typically the course is offered as a large lecture course. However, for many semesters one professor has offered a non-traditional section (HAL Online) for students preferring a problem-based format that may meet face-to-face occasionally during the semester but is taught mostly online. The course aims to develop scientific literacy through reading and online argumentation around real-world problem tasks, often presented with video cases. Students are assigned to small groups of 3-5 members that work online throughout the semester.

The units in the spring 2011 HAL Online offering that was the data source for this study were: I. Cognition and Culture; II. The Amazing Learning Brain; and III. Using Learning Science in Reflective Practice. Each unit comprised four or five weeklong lessons. During a lesson, students study multimedia resources about psychological science content drawn from textbooks and from video and news sources such as TED.com and The New York Times. In alternate weeks students either post a reflective personal blog that answers a problem-solving prompt, or participate in online collaborative problem-solving tasks. Blog posts and discussions are graded using a rubric that rewards understanding and intelligent use of course ideas.

Since 2007, the course has been a site for field testing Video Mosaic (Videomosaic.org), an NSF-funded research and development project that has created an online repository comprising an extensive, searchable, annotated collection of research video on children's mathematical reasoning and development. This collection, based on the work of researchers Robert B. Davis and Carolyn Maher (e.g., Maher, 2005) is a valuable resource that builds on extensive prior research including a longitudinal study following the same cohort of students through high school and beyond. Work described in this paper represents research and development with this valuable resource for teacher education.

The Problem

We hypothesized that students who argue better in our course will achieve more. Yet, despite a natural tendency to argue, students' arguments are often ill formed, lack evidence, and are incomplete (Kuhn, 2005). Moreover, individuals sometimes fail to understand what qualifies as evidence (Glassner & Schwarz, 2005). Struggles with constructing sound arguments are complicated by resource-rich environments resulting from advances in technology. Integrating evidence from a variety of multi-media resources has become an increasingly important component of arguing well. Our challenge is to design instruction that promotes students' development of sound arguments in resource-rich environments.

The approach favored by science educators (e.g., Cavagnetto, 2010) is immersion with scaffolding of argument. Yet this approach has practical limitations for our online university setting where there are pressures to increase enrollments despite lower instructional budgets. Scaffolding must be provided online to many students by a single faculty member unassisted, or with the help of relatively inexperienced teaching assistants who themselves may have poor argument skills. Recently universities have begun to offer relatively unsupervised massive open online courses (MOOCs) aimed at large-scale participation and open access via the WWW. Our aspiration is to serve large enrollments using an argument-based pedagogy. It is important, therefore, to investigate the viability of pragmatic alternatives to human guidance of argument online. One option is formally training students in argument *prior* to engaging them in pedagogies that require integration of conceptual content and evidence from multiple sources as a basis for well-reasoned claims.

Toward that end we conducted a controlled experiment within HAL Online to examine the effects, on individual learning and collaborative meaning-making, of a week-long unit offered early in the course that provided direct training intended to improve students' understanding of and ability to engage in good argument. The performance of a group receiving the treatment, *Trained Argumentation with Modest Scaffolding* (TAMS), was compared with an ecological control group that did not receive argument training: *Emergent Argumentation with Modest Scaffolding* (EAMS). Except for the treatment manipulation, both groups participated in an identical course of study and assessment. One hypothesis was that achievement for *individual* students, as measured by tests of comprehension and scientific literacy, would be higher for students who participated in TAMS. This hypothesis was strongly supported and is detailed in another analysis reported elsewhere (Gressick & Derry, 2013).

In contrast, the focus of analyses reported in this paper was on whether training in argumentation can influence aspects of *collaborative* meaning-making on tasks requiring small groups to integrate scientific ideas from course material with observations from real-world situations, to create evidential arguments for scientific explanations. Our analysis addressed the following research questions: Are there differences in how students in TAMS vs. EAMS use the scientific course material in their reasoning? Are there differences in how and how closely they attend to details of real-world cases provided by the problem? Is there evidence that argument training produces differences in how members of groups work together to blend their ideas and reasoning?

Theoretical Framework

The TAMS treatment was largely implemented thorough the lesson, *Adventures in Argument*, inspired by the Toulmin (1958) model that has served as the basis for many educational approaches using argumentation (e.g. Kuhn, 2005; Means & Voss, 1996 Stegmann, Weinberger, & Fischer, 2007). Toulmin's model is the basis of Halpern's (2002) *Analyzing Arguments*, a chapter in her award-winning text *Thought and Knowledge*, which was an assigned reading constituting a portion of the TAMS treatment. Her chapter emphasizes recognizing and using five components of good argument: conclusions, premises, counterarguments, qualifiers, and assumptions.

Arguing to support scientific explanations or theories is a social practice involving communication and persuasion. Berland and Reiser (2009) draw an epistemological distinction between the process of defending scientific explanations and the process of creating them, two key but distinct components of scientific practice (Kuhn, 2010). Although constructing explanations is central to scientific practice, it is not the only goal. Emphasis on constructing explanations can even undermine attention to evidence. Training students in argumentation, we hypothesized, would direct their attention to using course material as sources of evidence to justify and support explanations.

Data Source and Design

We conducted an *in vivo* experiment, which manipulates elements of instruction in a natural setting and observes the effects on student learning (e.g. Aleven & Koedinger 2002). The context of the study was the spring, 2011 offering of HAL Online. Forty-four students enrolled. A summary of the experimental design is shown as Figure 1. The treatment manipulation was the last lesson in the first course unit. The context for studying treatment effects on collaborative meaning-making were forum discussions in units occurring four weeks and eight weeks following the treatment. A separate Moodle course environment was created for each condition. These were identical except for the treatment-related manipulations.

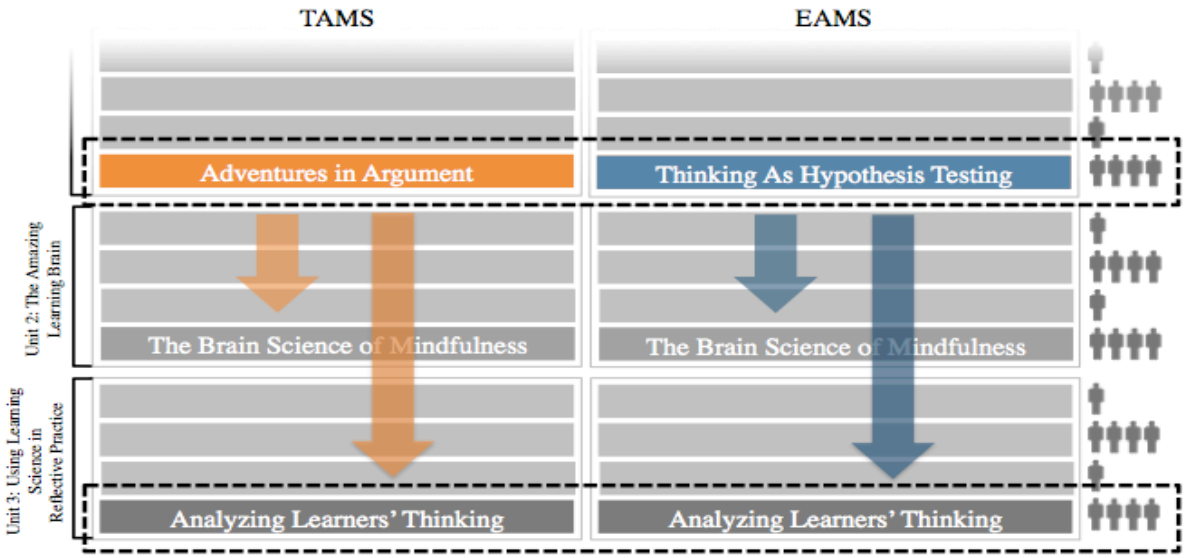


Figure 1. Overview of instructional design for TAMS and EAMS.

Using a within classroom nested design (Salden & Koedinger, 2009), students enrolled in HAL Online were assigned to small groups based on common interests as determined by self-report surveys. Small groups were randomly assigned to two conditions (described next). Groups comprised three or four students and, to avoid confounding the group dynamic, were maintained throughout the semester.

Treatment: Training in Argumentation with Modest Scaffolding (TAMS)

The goal of this lesson was to teach students the skill of making and recognizing strong arguments. Students read "Analyzing Arguments," a 50-page chapter on argumentation from *Thought and Knowledge* (Halpern, 2002). Following a quiz, students engaged in a collaborative forum discussion with their small group in which they practiced using ideas from reading to support analyzing and evaluating an argument in a speech.

Ecological control: Emergent Argumentation with Modest Scaffolding (EAMS)

In the EAMS control group students received an alternative weeklong lesson that focused on an alternate chapter of similar length and density from *Thought and Knowledge*, "Thinking as Hypothesis Testing." This chapter presented topics like the nature of variables, correlational versus experimental evidence, and using evidence for causal claims. During EAMS students completed a quiz and participated in a collaborative forum that employed the identical video speech but required designing a study to test the speaker's causal claims.

Description of Collaborative Forum Activities

Data were analyzed from two online collaborative forum activities during the semester. The first was titled *The Brain Science of Mindfulness.* This lesson occurred at the end of unit that directly followed the experimental argumentation (or ecological control) lesson (See Figure 1). Students read and viewed video about the scientific study of meditation practice. In their forum they debated the scientific merits of a proposed meditation-training program for a struggling middle school and were required to reach a group consensus.

The second forum analyzed, the primary focus of this paper, was in the lesson *Analyzing Learners' Thinking for Evidence of Preparation For Future Learning* and was a significantly more complex task that required sophisticated integration of multiple text and video sources. This lesson occurred as the final collaborative forum at the end of the course (see Figure 1). The primary goal was for HAL Online students to bring their knowledge of the claims of constructivist education together with claims about the nature of transfer proposed in a theory by Bransford and Schwartz (1999) to help them collaboratively examine elementary students' problem solving over time and in depth for the purpose of evaluating the scientific claims of the theory. This assignment represented a case of Berland and Reiser's "defending" or "persuasion" component of science practice, which they distinguish from constructing explanations. The explanations they were evaluating had been developed previously based on resources already encountered in Unit 3: *Rethinking Transfer* (Bransford & Schwartz, 1999) and *Should Schools Adopt a Constructivist Approach to Education?* (Windschitl & Hirsch, 2002). To seek evidence for these theoretical explanations, students were directed to access a series of six video clips from the Video Mosaic repository (Videomosaic.org), in which 11th grade students solved and justified solutions to a combinatorics problem, leading them to struggle with understanding Pascal's triangle and exponential reasoning. The 11th graders in the videos had been part of a cohort followed from early grades and

that had been immersed in constructivist learning environments through their years of schooling. HAL students had previously studied videos of this same cohort solving and justifying similar combinatorics problems within a constructivist educational setting in the 4th grade. The HAL Online students' discussion task was framed as follows and they were required to reach consensus.

> *What evidence do you find that early educational experiences have prepared the students in these video clips for future learning? How confident are you regarding claims that these students' earlier educational experiences have had an impact? What convinces you or would convince you?*

Method of Analysis

This study used analysis procedures for quantifying qualitative analyses recommended by Chi (1997) and outlined in the following stages.

First, forums were searched for instances where groups used concepts from the readings and evidence from the video series. As data were searched, a chain of reasoning (Chi, 1997) for each group was developed to eliminate multiple coding of the same concept within the same discussion thread. Because of the interconnected nature of collaborative meaning making, the group was viewed as the unit of analysis. However, connections to individual contributions are not lost in analysis and can be viewed on graphs in the results section of this paper.

Next, a sample of data was coded using the scheme in Table 1. The coding scheme follows phenomenon-based hypothesis coding (Miles & Hubermann, 1994; Saladña, 2009) that focused on the elements students leverage from course readings and the general ways that students applied these ideas to their discussions and, as a small group, engaged in meaning making. Our scheme was based on coding developed by Pena-Shaff & Nicholls (2004) and methods of collaborative meaning making described by Stahl (2006). Once the coding scheme was stabilized through discussion within our research group, reliability between two coders was calculated and reached 95% (Cohen's kappa .93) after two rounds of coding and discussion (Cohen, 1968).

Table 1. Description of coding scheme.

How do groups *integrate ideas* from text resources in discourse?		
Code	**Description**	**Example**
Quote text resource directly	Provides a direct quotation from an assigned reading	*"Information presented in the context of solving problems is more likely to be spontaneously utilized than … simple facts." (Rethinking Transfer)*
Restate text resource	Restates an idea from reading in own words	*The Bransford & Schwartz article mentions SPS testing… often fails to capture transfer…*
Extend concepts from text resources	Extends or applies a concept from reading to discussion	*The article on transfer also mentions the importance of this…for transfer to occur, the children cannot think of the concept in only one type of situation…*
How do groups *integrate ideas* from video of student learners in discourse?		
Quote video evidence directly	Provides a direct quotation from the video data	*Shelly says "my teacher's going to kill me because she knows I can do this."*
Restate video observations	Restates or summarizes observations from video	*[Robert] had already been sketching the tower problem from before this discussion was even underway.*
Extend video observations	Extends or interprets video data	*[Robert's] ideas are important to the group discussion and thinking in this clip.*
In what ways do groups *build* inter- and intra-subjective *understanding*?		
Up-take	Restates what another group member stated in prior post	*I think that the students definitely used their previous learning in these videos as well.*
Elaborate	Extends an idea previously mentioned by another group member in a previous post	*…I, too, thought about the concept "use it or lose it." In past readings about the aging brain, we read that the brain retains its plasticity…*
Draw personal connection	Makes a connection between content of discussion and personal experience	*They very quickly start with the triangle approach that we did in class and … related it to their previous experience.*

After coding, data were quantified and represented in a tabular format in order to find differences in patterns between treatment groups. Bar graphs were created to further reveal and enable study of patterns across groups and conditions. For each code we calculated group means, standard deviations, and treatment effect sizes. Our decision to focus on descriptive rather than inferential statistics was influenced by the relatively small number of groups in each condition.

Results and Discussion

Results of analyses from *The Brain Science of Mindfulness* (which did not require integration of video evidence) are briefly summarized as they replicate and strengthen findings. Our major focus is on the *Analyzing Learners' Thinking* forum. In both, TAMS clearly influenced student accountability to resources during meaning making.

Meaning-Making in the *Brain Science of Mindfulness* Forum

How Groups Used Resources

In both treatment conditions, the most consistently observed method of integrating ideas from the course was to extend the findings of research studies discussed in course readings to the case presented in the forum discussion task. The mean frequency of idea extensions for TAMS (M = 4.83, SD = 2.32) was higher than for EAMS (M = 3.0, SD = 1.67). The effect size for this analysis (d = 0.92) exceeded Cohen's (1988) guideline for large effect size (*d* = .80). Additionally, groups integrated ideas by directly quoting resources. Overall, groups in TAMS demonstrated a higher number of direct quotes in their discussions (M = 3.5, SD = 1.52) than groups in EAMS (M = 1.83, SD = 2.64), although the group with the most quotes was from EAMS. The effect size for this analysis (*d* = 0.80) was large. In TAMS, groups restated information from text resources from one to nine times over the course of the discussion, with 50% of groups demonstrating at least 5 instances of restating experts (M = 4.67, SD = 3.67). In EAMS, however, all groups produced five or fewer instances of restating experts (M = 3.17, SD = 1.17). The effect size (*d* = .62) was moderate.

How Groups Built Inter- and Intra-subjective Understanding

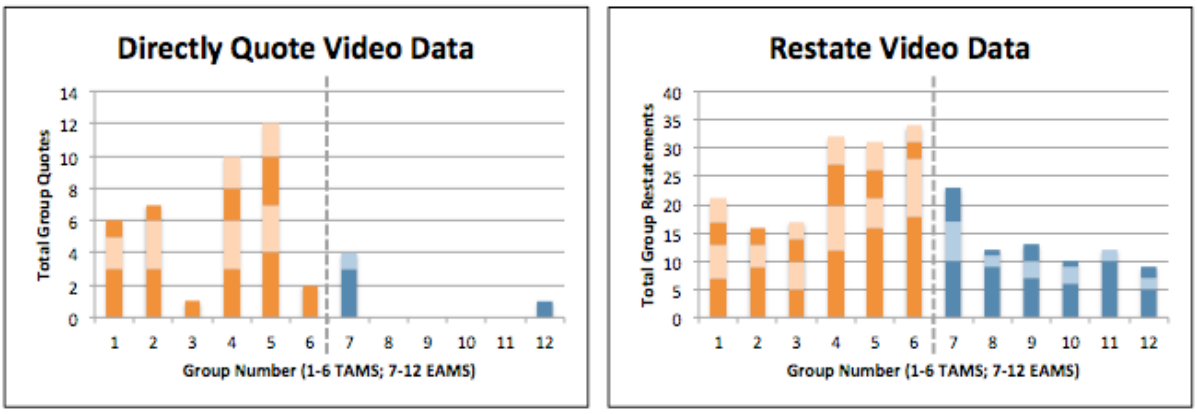
Groups in TAMS exhibited more up-take of ideas (M = 5.83, SD = 4.54) than EAMS (M = 3.17, SD = 2.23). The effect size for this analysis (*d* = .79) was high. Furthermore, groups in TAMS demonstrated more instances of elaboration of ideas (M = 2.33, SD = 2.07) than groups in EAMS (M = 0.5, SD = 0.84). The effect size for this analysis (*d* = 1.26) was large. In addition, all TAMS groups drew at least one personal connection, with most groups (67%) making at least 4 personal connections (M = 3.67, SD = 1.75). While 67% of groups in EAMS drew personal connections, only one group (Group 10) made more than one connection (M = 0.83, SD = 0.75). In most cases in TAMS but not EAMS, the personal connections made by group members were acknowledged and integrated into group discourse. The effect size for this analysis (*d* = 2.27) was large.

Meaning-Making in the *Analyzing Learners' Thinking* Forum

How Groups Used Resources

Groups in TAMS demonstrated a higher use of direct quotations from both text resources (M = 2.50, SD = 2.07) and from video data (M = 6.33, SD = 4.32) than groups in EAMS (text resources M = .1.17, SD = 1.6; video data M = .83, SD = 1.60). The effect sizes for both quotations of text resources (*d* = .72) and quotations of video data (*d* = 1.86) were large. Figure 2 provides a visual comparison of direct quotations from video data. Variations in color in each bar graph indicate contributions that were made by individual members of the group.

We examined how often groups restated (in their own words with conceptual correctness) ideas from text resources and video observations in their forum discussions. In both TAMS (M= 25.17, SD = 8.08) and EAMS (M=13.17, SD = 5.04), groups more often restated observations of the learners depicted in the videos than they did information from text resources (TAMS M=10.17, SD = 6.31; EAMS M = 4.0, SD = 1.26). Moreover, groups in TAMS for both types of resources made more restatements than in EAMS. The effect sizes for both text resource restatements (*d* = 1.63) and video restatements (*d* = 1.83) were large. Figure 3 provides a visual comparison of restatements compared across groups from video data.

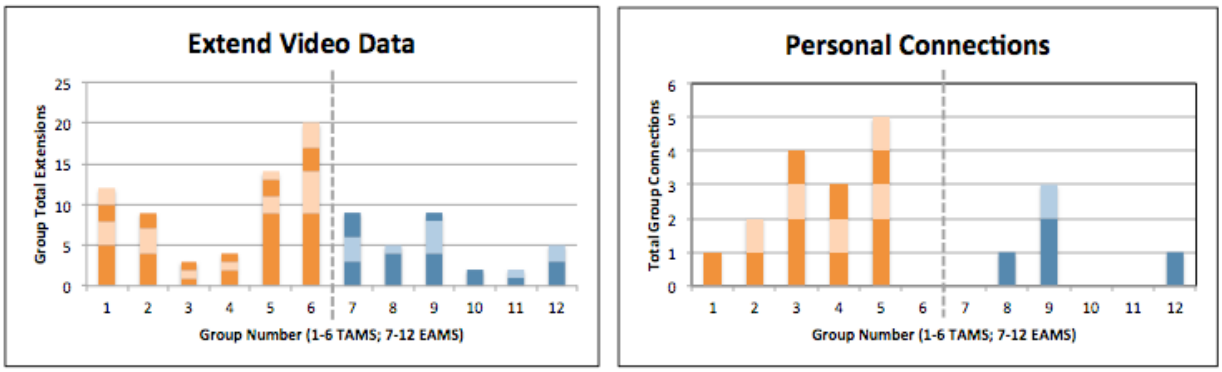


Figures 2 and 3. Direct quotations and restatement of video data resources in TAMS and EAMS.

We asked how resources were extended by groups as they constructed meaning in their discussion. Groups in TAMS were more likely to extend video (M=10.33, SD = 6.40) than EAMS (M = 5.33, SD = 3.14). Similarly, groups in TAMS (M = 7.33, SD = 4.84) were more likely to extend ideas from text resources than groups in EAMS (M = 4.0, SD = 2.0). The effect sizes for both text (d = .94) and video (d = .99) resources were large. Figure 4 provides a visual comparison of extensions of video data compared across groups.

How Groups Built Inter- and Intra-subjective Understanding

Similar to the findings for the *Brain Science of Mindfulness* forum, groups in TAMS demonstrated a higher average frequency of up-take of ideas (M = 8.50, SD = 7.42) than EAMS (M = 5.17, SD = 2.99). The effect size (*d* = .66) was moderate. Further, groups in TAMS demonstrated more instances of idea elaboration (M = 7.33, SD = 7.20) than groups in EAMS (M = 3.33, SD = 2.07). The effect size for this analysis (*d* = .88) was large. In addition, all but one TAMS groups drew at least one personal connection (M = 2.5, SD = 1.87). While 50% of groups in EAMS drew personal connections, only one group, Group 9, made more than one (M = 0.83, SD = 1.17). The effect size for this analysis (*d* = 1.10) was large (Figure 5).



Figures 4 and 5. Extensions of video data resources and personal connections in TAMS and EAMS.

An Illustrative Case of Precise Resource Use in TAMS

As presented in the data above, groups in TAMS demonstrated a more robust and precise use of resources in their discussions. The data presented in Table 2 exemplifies how precise, accountable use of data by a group's initial posting member served as a point of entry for other members of the group to participate in the data analysis and encouraged a collaborative meaning-making process. Subject 1 of Group 4 (TAMS) started the discussion with a specific set of "field notes" from his video study (see Table 2, post 2.1 below), including time stamps and direct quotes from the video clips. After this initial post, other group members may have modeled this first poster, adopting similar approaches in their own posts. In her response, Subject 43 offered an alternate, more abstracted organization of analysis focused on the participants from the video data (Table 2, post 2.3). What followed was a transformed analysis and synthesized summary by Subject 32 (Table 2, post 2.6).

Table 2. Abbreviated Group Forum Posts, TAMS Group 4

Post 2.1, Subject 1 (Initial Post)
I'm going to type my thoughts as I watch the videos to start off the discussion… ***Clip 1:*** The students are VERY aware that from somewhere, they have learned the tools to approach this type of problem. (Shelly says "my teacher's going to kill me because she knows I can do this")… ***Clip 2:*** Stephanie and Shelly are discussing how they are organizing the cases… ***Clip 3:*** I can't…understand what the teacher is asking…But at about 4:25 in this video … ***Clip 4:*** Although I'm finding it difficult to re-articulate Stephanie's explanation for why the Pascal's triangle explains the pizza situation, it's sounding pretty logical and convincing.. ***Clip 5:*** Stephanie shows an indication of transfer right off the bat: "we worked on it in 8th grade"… ***Clip 6:*** Amy gets a little bit more engaged in this one… At about 9:20 in the final clip I finally became aware of why there is a "2" as the base in the exponential expression describing the pizza problem…
Post 2.3, Subject 43: Okay so I started out taking notes on each clip like [subject 22] and [subject 1], but I came to basically the same conclusions…so..I'm…gonna add some observations about the group and each kid in it....
Post 2.6, Subject 32 (Group Summary)… drawing a connection between the Tower problem and Pascal's Triangle allowed first Robert and then the rest…to better understand the problem, both through abstract verbal reasoning and spatial imaging, and…the students use the abstract model to both…justify answers to previously solved problems and reconcile [current] confusion, which is a sophisticated and enduring form of knowledge…

In contrast, Subject 42 of Group 11 (EAMS), initially posted an abstracted, summative analysis of the data, organized around themes (see Table 3). Unlike the approach taken by Subject 1 in Group 4, Subject 42's original post relied heavily on her restatement and summary of observations from the video. This initial post lacked the accountability and precision with which Subject 1 leveraged the video resources. While there are some positive qualities of Subject 42's post, other members of the group did not engage in analysis as actively or in-depth as was observed across members of Group 4. One suggestion of why this might be is the degree to which Subject 42 had abstracted her observations to support the claims in the text. Because of this, a clear model of her meaning-making process – a point of entry into the collaborative process – was not provided to the other members of the group, as it was in Group 4. What had resulted from the approach adapted by Group 4 was a rich, integrated understanding of how specific evidence from the video could be used to promote an argument in support of the preparation for future learning theory of transfer. This example suggests the importance of precise resource use as a means to facilitate collaborative meaning-making processes.

Table 3. Abbreviated Initial Group Forum Posts, EAMS Group 11

Post 3.1, Subject 42 (Initial Post)
I think these students' prior experience with the constructivist approach…has prepared them for future learning. ***Transfer***…They used algebra that they had learned prior to this lesson to help them solve the problem…This shows that the students were able to transfer information they had previously learned…they continued to double check it and try it from other angles. This shows constructive transfer because… **Expert**… Stephanie proved that she had a deep understanding of the problem because she was able to understand and recognize almost immediately that the way the other girl had done the problem was correct even thought it was different than her way. This shows Stephanie exhibiting skills that an expert would exhibit. **Mix**…In the very beginning the students Stephanie and other girl (don't remember her name) discussed how they were going to go about solving this problem…This also shows that their knowledge was transferred… organizing thoughts and looking for deeper concepts than just a formula show skills of an expert.

Conclusions and Scholarly Significance

Discussions in groups that received argument training were consistently more enriched by references to and sharing of the course material. In the more complex problem that involved analyzing videos of children's problem solving over time to determine the credibility of a scientific theory, students trained in argument correctly incorporated into their discussions more scientific material, and they conducted more exacting and careful search of videos to identify evidence related to the theory. A formal course of training in argument might well result in more accountable discussions, where groups integrate more from resources into their discussions. That they use more direct quotes, for example, suggests attention to preserving the words of credible sources and may indicate a heightened precision regarding data from sources. The findings of this study, which demonstrate effects on student resource use occurring many weeks after *Adventures in Argument*, indicate that training prepared students for future learning (Bransford & Schwartz, 1999). The study contributes to a discussion on how to optimally approach argumentation as pedagogy and provides support for the direct training approach, at least for college-level learners in online environments that employ argumentative pedagogy. Moreover, this approach offers a viable alternative to more resource-intensive immersive approaches for online environments. While the unit of analysis in this study was the small group, data for individual students was visually supplied. The patterns of individual involvement open an area that requires further investigation: although many groups experienced participation from multiple members, there was room for improvement. However, combined with a companion study (Gressick & Derry, 2013) showing positive effects of training on individual student learning, this research shows that direct training in argumentation is a promising intervention.

References

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26, 147-179.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* , *24*, 61-101. Washington DC: American Educational Research Association.

Berland, L. & Reiser, B. (2009). Making sense of argumentation and explanation. *Science Educatio*n, 93, 26-55.

Cavagnetto, A.R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research*, 80(3), 336-371.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *Journal of the Learning Sciences, 6*(3), 271–315.

Cohen, J. (1968). Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70 (4): 213–220. doi:10.1037/h0026256. PMID 19673146.
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum.
Glassner, A. & Schwarz, B. B. (2005). The antilogos ability to evaluate information supporting arguments. *Learning and Instruction, 15,* 353-375.
Gressick, J. & Derry, S.J. (2013). The Influence of Training in Argumentation on Students' Individual Learning Outcomes. In Rummel, N., Kapur, M., Nathan, M, & Puntambekar, S. (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings* (Vol. 2, pp. 38 - 41). International Society of the Learning Sciences.
Halpern, D. (2002). Thought and Knowledge (4th ed.). Mahwah, NJ, Erlbaum.
Hay, D.F., & Ross, H.S. (1982). The social nature of early conflict. *Child Development*, 53, 105-113.
Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, 810-824.
Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.
Larson, A.A., Britt, M. A., & Kurby, C. (2009). Improving students' evaluation of informal arguments. *Journal of Experimental Education, 77*(4), 339-365.
Maher, C.A. (2005). How students structure their investigations and learn mathematics: Insights from a longitudinal study. *Journal of Mathematical Behavior,* 24, 1-14.
Means, M. L. & Voss, J.F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction,* 14, 139-178.
Miles, M.A. & Hubermann, A.M. (1994). Qualitative data analysis (2nd ed.). Thousand Oaks, CA: Sage.
Noroozi, O., Weinberger, A., Biemans, H.J.A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review*, 7, 79-106.
Pena-Shaff, J & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, 43 (3).
Saladaña, J. (2009). *The Coding Manual for Qualitative Researchers*. Los Angeles: Sage Publications.
Salden, R. J. C. M. & Koedinger, K. R. (2009). In vivo experimentation on worked examples across domains. Symposium at the Thirteenth Biennial Conference of the European Association for Research on Learning and Instruction.
Sandoval, W. (2005). Understanding students' practical epistemologies and their influence. *Science Education*, 89, 634-656.
Schworm, S. & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99, 285-296.
Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. MIT Press.
Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2, 421-447.
Toulmin, S. (1958). *The uses of argument.* Cambridge: Cambridge University Press.
Toulmin, S. (1972). *Human Understanding, Volume 1: The Collective Use and Development of Concepts.* Princeton: Princeton University Press.
Windschitl, M. & Hirsch, E.D. (2002). Should Schools Adopt a Constructivist Approach to Education? In Abbeduto, L. (Ed.). *Taking sides: Clashing views on controversial issues in educational psychology*, 2nd edition. Guilford, CT: Dushkin/McGraw-Hill.

Acknowledgments

This material is based upon work partially supported by the National Science Foundation under Grant No. 0822189. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Collective Engagement in a Technologically Mediated Science Learning Experience: A Case Study in a Botanical Garden

Fariha H. Salman, Heather T. Zimmerman, Susan M. Land, Penn State University, University Park PA 16802
Email: fxh139@psu.edu, heather@psu.edu, sland@psu.edu

Abstract: This paper presents an interactional case study from *Tree Investigators*, a research study designed as a technologically mediated tour of an arboretum where children aged 7-11 collaboratively learn about the characteristic features of different types of trees. Throughout the tour, children are facilitated by a Naturalist and use mobile technology (e.g., iPads) to focus on specific characteristics of trees on their touch screens while observing the trees and discussing about them. This analysis focuses on a group activity where children use a mobile app to identify a *mystery tree*, analyzed through video-based Interaction Analysis. The findings reveal a collective engagement afforded by a coordinated interaction between sensory modes (verbal, gaze, touch, spatial) and mobile technologies (iPads, AR content). The purpose of this analysis is to help researchers and educators utilize the analytical concept of the *collective* when designing or examining mobile learning activities outside of school.

The *Tree Investigators* research and design project investigates science learning within technologically enhanced outdoor informal learning institutions (ILIs) such as nature centers and arboretums. *Tree Investigators* includes a field tour at an arboretum where families and children are facilitated by naturalists to collaboratively learn about the characteristic features of different types of trees. Throughout the tour activity, children are accompanied by their families and are encouraged to use mobile technology (e.g., iPads) to discuss the scientific characteristics of trees as they observe trees. The analysis in this paper focuses on the coordination of sensory observations, interactions, technology, and science content when children were facilitated to use a mobile application to identify an unknown tree, called the *mystery tree*. This paper contributes an account of collective engagement, coupled with learners' sensory interactions, as a theoretical tool that can aid in the research and design of mobile computing to support out-of-school learning in ILIs.

Conceptual Framework

This paper focuses on interactions between learners, mobile computers, and an outdoor learning center to exemplify a collaborative educational design which includes learner-centered pedagogy relevant to learning outside of school. With this focus, our work draws on two theoretical literatures: *collective* engagement (Thomas & Brown, 2011) and learners' interactions with technologies. The structure of our mobile learning activity includes learner-centered, small group engagement facilitated by a Naturalist where learners work on authentic problems to acquire new scientific information as part of a visit to the Arboretum at Penn State.

The learners' activities occur within the context of emerging technologies, which present a potential for analyzing what Thomas and Brown (2011) call the *collective*, which is a highly collaborative problem solving system relying on the complex, real time coordination of various resources — including people, skills, technologies, and interactions. The concept of the *collective* relates to the framework of distributed intelligence (Pea, 1993, White & Pea, 2011) whereby intelligence is seen as spread across social and material resources. In the same vein, "collective intelligence" is enlisted amongst the eleven core social and educational skills for children in the emerging participatory culture of the future (MacArthur Foundation, 2006, p. 4). Also, various attempts at re-imagining education, such as the Centre for Educational Research and Innovation (e.g., '*Schooling for Tomorrow'* and '*Future of Higher Education*'), direct educators towards a future where learning will require cross-disciplinary expertise encompassing multiple ways of knowing. Concepts like distributed intelligence and the *collective* enable researchers to understand that these multiple ways of knowing do not reside within the individual; rather, knowing locates itself in the dynamics of coordinated interactions of human and technological resources. An important characteristic of the *collective* as emphasized by Thomas and Brown (2011) is that the collective is "defined by an active engagement with the process of learning" (p. 52) while "providing access to an increasing number of resources managed by a technological infrastructure" (p. 53) since it is "well designed to facilitate peer-to peer learning, their raison d'être" (p.53). We rephrase these characteristics as: (1) active engagement with the learning process and (2) accessing multiple human and technological resources.

It is important to note that collective engagement invokes the concept of multimodality or multiple semiotic modes (Hodge & Kress, 1988; Kress, 2005; Kress, et al., 2001; Lemke, 2002) from the field of social semiotics. Kress et al. (2001) explains multiple modes within learning "when learners actively engage with all modes as a complex activity in which speech or writing are involved among a number of modes" (Kress et al., 2001, p.1). The importance of multiple modes in learning specifically science literacy is furthered by Lemke

(2002) who asserts that all meaning resides in the integration of complex material systems that span across temporal, spatial scales which can be seen as semiotic resource systems—separable only analytically. One semiotic resource is digital technology (Baldry & Thibault, 2006) that combines and unfolds other semiotic resources in new and innovative ways (O'Halloran, 2009). Researchers (e.g., Mann & Reimann, 2007) also emphasize the role of mobile learning technologies as mediating tools—acting as a cultural intermediary between the learner and his or her social and physical environment. In keeping with this view, we understand that mobile technology serves as a semiotic resource for families' meaning making.

Studies (e.g., Kahr- Højland, 2011; Sung et al., 2010) where out-of-school learning activities utilized mobile technologies have recommendations that align with characteristics of the *collective*. Sung et al. (2010) in their museum-based study found that students using a mobile problem solving guide system fared well in terms of interactions and learning-related discussions. Kahr- Højland's (2011) findings from a science centre study favor a narrative-based exploratory design for meaningful technological scaffolding. However, both studies lament learners' disengagement with the exhibits, and a lack of deeper analysis, especially when utilizing technologies. They recommend a detail-oriented design that explicitly directs learner's attention to some specific aspect of the materiality (e.g., texture, actual size) on-site in order to engage learners in in-depth discussions and careful study of the exhibits, instead of a focus on just the technology (Hsi, 2003).

Research Question

How do mobile computers interact with sensory semiotic modes to support collective engagement of learners while exploring trees in an ILI?

Methodology

Setting and Participants

The study is set at the Arboretum at Penn State, a botanical garden that displays trees from across the world. The Arboretum at Penn State is a 370-acre ILI; it features 35 acres of groomed gardens, a children's garden focusing on Pennsylvania natural history, and an old growth stand of hardwood trees with walking trails. This ILI offers an outdoor space that could be designed to enable a participatory, immersive science learning experience. This aligned with the *Tree Investigators'* intent to focus on informal spaces where families can enjoy and learn from in situ scientific phenomena related to trees.

Ten families were recruited for the *Tree Investigators;* they were members of a nature center close to the Arboretum. The families were strategically selected: (a) as heavy users of ILIs and (b) because of their affinity for outdoor experience of life sciences. Our strategic selection of ILI visitors is aligned to research practices commonly used in empirical museum studies (e.g., Allen, 2004; Leinhart, Crowley & Knutson, 2002).

The *Tree Investigators* project uses Augmented Reality (AR) to bring web-based media to a smart phone or tablet such as an iPad. In designing the mobile website (Zimmerman, Land, McClain, Mohney, Choi, & Salman, 2013), we aligned with recommendations for mobile computers on (a) the importance of personalization to the learners' agendas (Kearney, et al., 2012), (b) brief just-in-time interaction with the device, to facilitate learning through conversation (Hsi, 2003; Kahr-Højland, 2011), and (c) matching the learners' expectations of the experience to the affordances of the device (Looi, et al., 2010; Sung et al., 2010). After interacting with the mobile website, learners were presented with a cumulating task to identify a mystery tree.

Data Collection Techniques

Data was collected in the Fall of 2011 with 10 families (25 participants) including 15 children aged 7 to 11. Data collection occurred on weekdays when the schools were closed for teacher professional development. The research team organized families in smaller groups, which were each facilitated by a Naturalist during the 60-minute tour for each group. Videos of each of the 5 field tours were recorded, resulting in 5 hours of recorded data. The video data were transcribed into approximately 150 single-spaced pages of text.

Data Analysis Procedures

In accordance with our research question, a video-based Interaction Analysis was conducted (Derry et al., 2010; Heath & Hindmarsh, 2002; Jordan & Henderson, 1995) to explain the interaction between the children's sensory modes and mobile technological modes (e.g., apps & AR content) in identifying a mystery tree as a collective engagement enterprise. Interaction Analysis aligned with the research question as it seeks to "investigate human activities such as talk, nonverbal interaction and the use of artifacts and technologies identifying routine practices and problems and the resources for their solutions." (Derry et al., 2010, p.1). Also, interaction analysis is best suited to ethnographic approaches (Heath & Hindmarsh, 2002), and the video-recorded field tour of the *Tree Investigators* followed ethnographic methods. These analytical procedures allow for capturing the complexity of interaction in its various steps including: video review sessions, cannibalizing, transcription, selecting events, and extracting fragments; these are described below.

Data analysis involved reorganizing the video-based transcripts into multimodal transcripts (Kress et al., 2001), which included the descriptive dimensions of the sensory modes like gaze, speech, and body posture and positioning. Reorganizing the transcripts to highlight both verbal and nonverbal elements allowed the researchers to make analytical insights related to collective engagement. Both verbal and nonverbal aspects were recorded using time as an anchor. Another criterion of data reorganization was preserving the interactional sequence in which sensory modes appeared (e.g., if the action preceded the verbal or if the mobile app was being used in conjunction with a sensory mode such as talk or touch).

The criteria for recognizing instances as *events* according to Jordan & Henderson (1995) are "coherence in some manner" along with "official beginnings and ending" (p. 20). Heath, Knoblauch and Luff (2000) add an additional criterion that events allow for the "enabling retrieval of critical information" (p. 313). Using these principles, the learners' engagement with the mystery tree task episodes each formed one event. The events selected for analysis in this paper were: (a) related to collective engagement around the mystery tree task episodes, which totaled approximately 100 minutes of video, and (b) focused on the video sections where the interaction between sensory and technological modes were visible. The segments were identified based on the problem solving event of recognizing six scientifically relevant tree features (i.e., branching patterns, leaves/needles [shape, arrangement, margin/edge], fruit elements [color], flowers [shape, color]); these six scientifically relevant tree identifiers were included on the *Tree Investigator* mobile website and app. The *US Trees* app was used during the mystery tree task which has an interface that presents the identifying features of trees to support identification. *US Trees* is designed so that the user sees multiple possible options for each feature and can select one option after observing that feature on a real, virtual tree or an image of the tree. The selected options are then configured as possible tree identities in the app's database.

This paper presents one analysis of one mystery tree task event, within which seven segments are presented to explain the interaction between the learners' sensory modes (e.g., gaze, touch, verbal, spatial) and the available mobile technologies (mobile app, AR content) in the process of collective engagement to identify a tree. The seven segments are those identifying features that are noticed by the group in this particular event; the seventh segment is where the results are discussed. Moreover, the order in which these segments appear reflects the participants' learning agenda during their collective engagement.

Data and Findings

From the analysis of the data, we found the two characteristics of the *collective* (Thomas & Brown, 2011) in the interactions of the children equipped with both sensory and mobile technological resources in the informal education design of the mystery tree activity. We first analyze the interactions within each of the seven segments while referring to the data transcripts. Next, we discuss the event as a whole, guided by the two characteristics of the collective: (1) active engagement with the learning process, and (2) accessing multiple human and technological resources.

The collective engagement event using the mobile computer began when the three children, Lydia (aged 8), Emmy (aged 8), and Greg (aged 11), were introduced to the *US Trees* app to identify a tree's scientifically relevant features as a means to identify an unknown tree (i.e., mystery tree). The Naturalist only intervened where the children needed facilitation. In the first two segments (below), Greg expressed curiosity about the flower (line 867) whereby the Naturalist used the AR content to show a digital photograph of the springtime flower (lines 867-893) and fruit (lines 897-911), to the learners.

867. [00:39:17.21] Greg: How do you figure out the flower?

868. Naturalist: Right. So when you're ready for the flower, you can come over here- or the fruit-
869. both the fruit and the flower I'm gonna have to show you using this special thing.

870. Lydia: (walks over and sits down next to Naturalist & Greg where a laptop and AR tags
871. are laying) This is what I think.

872. [00:39:33.23] Naturalist: Soo. (picks up large AR tag) For the flower- I'm- right now what
873. I'm going to do is augment the flower and the fruit for you. And by augmenting I mean I'm
874. gonna use a digital representation of, um, the fruit and the flower to show you so that
875. you can- because right now (points up at tree), it's the fall and so we can't see those two
876. features and so I have to show you using this.

---

900. Naturalist: So there's the color of that fruit! (holding AR tag in front of laptop)

901. [00:41:03.09] Lydia: (taps iPad screen) There!

902. Naturalist: Did you get it? Everybody got it?

903. Rachel: You got the fruit, Emmy?

904. Emmy: Yeah, I think so.

905. [00:41:08.14] Naturalist: Alright. Great. So let's see- I think we're all good with that. (shuts
906. laptop lid) I'm going to shut this up so that it doesn't get- whatever, dirty (laughs)

907. Lydia: And this right here! What about this right here? (point iPad screen towards Naturalist
908. and pointing with finger at a place on the screen)

909. Naturalist: Right. So that is the flower shape, do you think- which one of those three do you
910. think that flower looked like?

911. Lydia: (taps screen) Right here.

The third segment (below) includes a conflict that arose between the children due to different possible trees showing up as possible mystery tree solutions on the *US Trees* app, which meant that there was some confusion in identifying the tree's salient features. As a result, the Naturalist suggested the three children to "work together to see what we selected" (line 919) and guided them to look at the 'leaf shape' (line 920) in the fourth segment. The video transcript shows that the children reached a consensus about the leaf shape, which the Naturalist reinforced (lines 921-923). The fifth segment shows a transition towards the branching structure (lines 929-932). The children chose 'opposite' on the *US Trees* app after a sensory engagement with the tree on-site and after they confirmed this observation with each other. For example, Lydia, moved closer to Emmy to look at her iPad screen and asked "What did you do?" (line 928), Figure 1. They touched and unfolded the leaves to observe and feel the branching pattern as shown in Figure 2. Even before the Naturalist asked them to pull down a "branch with leaves to take a closer look" (lines 923-924), Figure 3 shows the children attended to the branching structure by moving closer to the tree, by holding their gaze on the leaf pattern.

919. Naturalist: //let's all sort of work together to see what we selected (walks to and stands by tree).

920. [00:41:54.16] Naturalist: Let's look first at the leaf shape (3 youth gather around tree with
921. Naturalist). What did, um, what did- let's see (leans over to look at Greg's iPad screen). Okay!

922. Lydia: Right here? (Naturalist looks at Emmy's screen then Lydia's)

923. Naturalist: Uh huh. Yeah! I think that, I think yer- hittin' the nail right on the head (pulls
924. down branch with leaves for youth to take a closer look), cuz if you kind of unfold it, it
925. looks- it looks like that one! (points at Lydia's screen)

926. Lydia: Mm hmm!

927. [00:42:13.15] Naturalist: I think that's right, okay!

928. Lydia: (to Emmy) What did you do? (leans over to look at Emmy's screen)

929. Naturalist: Soo. (looks at Greg's screen) And everyone picked the opposite branch- yeah!
930. Wow! Good job you guys! Okay, so, so you can see here that it is in fact opposite branching,
931. (holding bare branch out and pointing towards the node) those two branches came right
932. out of the same…point, right there. What's the next one? (leans over to look at Emmy's
933. screen) What's the next one?

934. [00:42:33.02] Lydia: Over here?



Figure 1



Figure 2



Figure 3

Greg, Lydia, and Emmy used the *US Trees* app on the iPads together to support their observation and identification practices. Once the children verbalized their identification, the Naturalist also pointed at the trees' nodes to draw attention to the 3-dimensional, actual tree specimen. Here, the learners coordinated their interactions between the sensory modes (i.e., touch, observation) and the mobile technology, as they moved back and forth between them. The learners' sensory experience involved gaze, gesture, touch, and talk, which helped the children to identify and confirm the branching structure. The mobile app presented the children with specific options of branching structure, which in turn focused their attention on the scientific aspects of the mystery tree. The constraint of the app channeled their noticings (Kellah, 2010; Yew & Schmidt, 2012) to the scientific knowledge that is relevant in this learning moment—in keeping with mobile technology pedagogy that enables domain-specific thinking strategies and knowledge. The three children interacted with the dynamic *screen* that empowered them to navigate at their own pace and interest (Thomas & Brown, 2011) and also gave them the relevant results at the end of their engagement. This app's process of dynamic assessment differs from

picking identity markers on the static paper (as in a book) and is an important learner-centered attribute of a mobile computing environment for ILIs.

The sixth segment (below) began with confusion about the leaf margin (i.e., the texture of the edge of the leaf). Unlike the fifth segment above where there was a consensus about the branching structure as 'opposite', two of the three children, Emmy and Lydia, were confused about the 'look' of the leaf margin. Since the field tour is happening during autumn, the older autumn leaves have a different shape and texture than younger summer leaves, which the *US Trees* app featured. The children's sensory experience told them that the leaves were "kind of bumpy" (line 936) and the children confirmed their observation by alternating between feeling the leaves and rereading the options on the mobile app. The Naturalist encouraged their coordination of these sensory and technological modes, and she only confirmed Greg's selection (line 939) when he approached her to show that his *US Trees* identity "match" (line 938). The Naturalist knew that the 3-dimensional on-site leaf had a time restriction that called for manipulation of a fourth dimension (i.e., time of year). She directed the two confused children to consider a different 'look' of the older autumn leaves when she said that "it's gotten crinkly and they are not as green anymore" (line 944). She attended to this instance, connecting the 'now look' of the autumn leaves with the 'before look' of the summer leaves. She said: "I think it's crinkled up and so it looks like there's a- there's a tooth to it but I actually think that it's probably smooth when the leaves- actually I can augment that too" (line 946). At this point the augmented image of the leaf is shown to the children so that they have a technologically accessed experience of the summer leaf.

936. Emmy: I think it's kind of bumpy (reaches up to touch a leaf on the tree)

937. Naturalist: You think it's kind of bumpy...

938. Greg: I have one match (tilts iPad screen towards LRM)

939. [00:42:43.13] Naturalist: Ahh! I think (smiles and nods at Greg) you might be on the right
940. track.//

941. Lydia: //It's kinda... (touching a leaf)

942. [00:42:48.24] Naturalist: You know what? I think that, it's actually pretty smooth, I know- you
943. guys do have a very sharp eye looking at that, because it does, a little bit, because it's gotten
944. crinkly, and they're not as green anymore, I think it's crinkled up and so it looks like there's
945. a- there's a tooth to it, but I actually think that it's probably smooth when the leaves-
946. actually, I can augment that, too. (Rachel laughs off camera)

947. [00:43:12.07] Naturalist: Lemme show you the green leaves (walks back to laptop on ground),
948. so we can- you don't have to take my word for it.

949. Rachel: It's a little bit harder when they're old and die.

950. Naturalist: Exactly (walks back to group carrying bag with AR tags)

951. [00:43:25.23] Naturalist: (opens laptop lid) Alright.

952. [00:43:42.02] Naturalist: Hellooo? (moving AR tag around in front of laptop).

953. Rachel: oohh.

954. [00:43:50.05] Naturalist: So, what do you think?

955. Lydia: I'm riiight. (taps iPad screen)

956. Naturalist: Oop!

957. Lydia: (sits up and looks at LRM) This.

958. [00:43:55.11] Naturalist: (singing) Technology. There we go.

959. [00:43:58.14] Naturalist: So whaddya think? Greg, what did you pick?

960. Greg: I picked oval.

The transitions in the above segments illustrate how mobile technology afforded a rich layering of realities, by varying time. The support in coordinating sensory and technological modes and using augmented reality to show variation over time, all contributed towards the collective engagement at the Arboretum. Particularly, in the seventh segment (lines 947-960), the use of AR, coupled with on-site observations and information from the app where the children themselves compare and self-assess their own identification as when Lydia exclaims "I'm riiight" (line 955) and Greg identifies the leaf shape as "oval" (line 960).

Discussion

The discussion section is organized by the two defining characteristics of the *collective* (Thomas & Brown, 2011) to examine the interaction between mobile technologies and sensory semiotic modes to enable collective engagement of learners within an informal education design. We focus on two key areas of theory: (1) active engagement with the learning process and (2) accessing multiple human and technological resources.

Active engagement with the learning process

The collective engagement of the three learners meant that they each took on an active role as the problem of identifying the mystery tree unfolded naturally (Kellah, 2010). The learners' interactions in the mystery tree activity reflect the individuals' learning agendas, which were encouraged by the learner-centered design of the integration of the Naturalist's pedagogy with the content from the mobile computing environment. Each learner accessed the sensory and technological semiotic resources. Sometimes the children worked individually while next confirmed with their group members (line 928) while other times, they attended to their group members by looking at what they were doing but picked their own option without confirmation from the peers (line 938). Throughout the activity, each learner exhibited an active engagement not only in identifying the mystery tree, rather in making sense of the identification problem by "riddling one's way through the mystery" (Thomas & Brown, 2010: p.9). This multimodal, collective sense making was seen when Emmy and Lydia were confused about the 'look' of the autumn leaf margin calling it "kind of bumpy" (line 936). They did not confirm the *US Trees* selection of the leaf margin until they were sure—by comparing the feel of the leaves to the options on the *US Trees* mobile app. The three children also involved the Naturalist in facilitating their thinking about this particular tree's feature. Incorporating supports from the adults on-site showed how learners engaged actively by utilizing all available resources to identify the mystery tree according to their personal learning agenda, which Thomas and Brown (2010) refer to as tinkering. Within the design of the activity, the Naturalist acted as a facilitator guiding them with their thinking process by her talk or by introducing AR content as technological semiotic resources (lines 942-946). This tinkering support becomes most visible when Greg expressed curiosity about the flower (line 867) since the tree had no flowers as it was autumn; here, the Naturalist used the AR content to access a digital photograph of the springtime flower (lines 867-876) and fruit (lines 900-911).

Moreover, the learners' active engagement was afforded by the structure of the mystery tree activity which takes a manageable form as it presented (yet limited) the available identifying features: branching patterns, leaves/needles (shape, arrangement, margin), fruit elements (color), flowers (shape, color) by means of the mobile app. Also, the repeated use of the same six features across modes supported the learners' sense-making since the children had already encountered these ideas in the earlier part of the Arboretum. Here, the technology afforded by the *US Trees* app aids in structuring the problem such that learners could access it in its composite and manageable parts that were linked to the children's prior knowledge. In keeping with recommendations (Kahr- Højland, 2011; Sung et al., 2010) that utilized mobile technologies in engaging activities, this channeling directed the learners' attention to specific features of the tree's materiality (e.g., leaf texture, color, shape.) so as to afford both engagement with and identification of the tree. Also, this repeated six feature structure functioned as a pedagogical tool whereby one iPad screen became the reference point for all participants (including the Naturalist) when converging together from where they depart into their own process of mystery-solving. Interestingly, this occurred at points of transition from one tree feature to the next- example: line 934 when Lydia asked: "over here?" and all attended to her iPad screen. This was in response to the Naturalist's directive of: "So, what's the next one?" (line 933) after they identified the branching pattern.

Accessing multiple human and technological resources

The interaction between the available semiotic resources at the Arboretum allowed the three learners and the Naturalist to move back and forth between the sensory modes and technological modes. Throughout the learners' participation in solving the tree identification mystery, they do not appear to be conscious of the resources used in their collective engagement—especially their reliance on their own sensory modes when observing the trees. This overlooking of their (learners) sensory modes was also revealed in the post-tour interviews where when the children reflected on their experiences mentioned only the device as the tool helping them identify the mystery tree. For example, one boy reported that "I could never figure that out without looking all around that app on the iPad". Across all the learners, the children only noticed the role of technological modes in this activity. Lemke (2002) asserts that we never make meaning with only the resources of one semiotic system and we glide through the various semiotic systems in "relatively automated ways" (p.2) and this study reaffirms his claim that learners use various systems but perhaps without notice. Figures 4-7 above capture this pattern: the children first looked up the features on the app [see gaze onscreen, Figure 4], then they moved closer to the tree to look at a particular feature [see spatial move and gaze on on-site tree, Figure 4 & Figure 5], next they looked at the app options [see gaze onscreen, Figure 5] followed by touching and feeling that feature on the tree specimen multiple times at the Arboretum. When the children were confused about the leaf margin, each child not only touched one leaf multiple times but the children also touched and

observed multiple leaves on the tree for confirmation [i.e., touching concreteness and confirmation through sensory modes in Figure 6]. After this sensory confirmation from the leaves at the Arboretum, the children picked one option on the app [i.e., technological mode/app]. This process of engagement followed further confirming and getting to know what their peers have selected [see Figure 1] as illustrated when the children who were not sharing an iPad were seen to go to at least confirm and check their selection such as in line 928 when Lydia asked Emmy; "what did you do?" This confirming of specimens on-site and on others' app choices was a more frequent move when a tree's identifying features had a seasonal or time restriction (e.g., deciduous leaves, flowers, fruits).

Kress's (2005, 2001) and Lemke's (2002) perspectives explain these children's engagement patterns as the inter-semiotic processes through which semiotic choices integrate to create meaning wherein the learner is empowered to make choices of which semiotic mode to use. The children also approached the Naturalist who used augmented reality (AR) to bring certain tree features to the children to technologically enhance the learning experience [i.e., technological mode/AR, see Figure 7]. The AR materials brought a digital 'summer time' to the Arboretum to be superimposed on the 'autumn time'. Bringing together multiple seasons allowed for a complete identification of the mystery tree. In summary, throughout their experience at the Arboretum we have shown that these three children moved smoothly back and forth from sensory to technological modes to access needed information to identify the tree. In the context of this study's findings, we posit this as learner's choice of representational modes within the framework of the collective engagement.

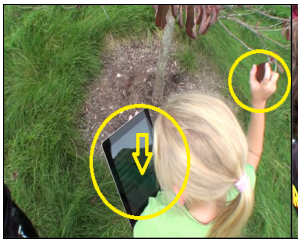
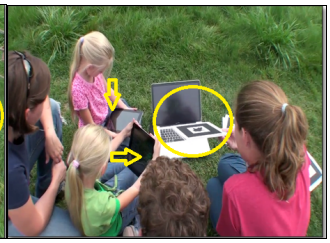

Figure 4


Figure 5


Figure 6


Figure 7

Conclusion

This exploratory study analyzed the interactions between various semiotic resources in the context of an informal educational project at the Arboretum at Penn State. It shows how learners move across resources and modes to develop their own paths towards collectively identifying a mystery tree. The collective (Thomas &Brown, 2011; 2010) was much more than just three children working together; it included the children making decisions about what was included in that collective (i.e., which resources and how to use them) while tinkering through the problem. Our findings, guided by the defining characteristics of the 'collective', include learners' active engagement in utilizing human and technological resources. Our interaction analysis of modal transitions revealed the visible paths and choices of semiotic resources that these three learners picked to solve the mystery tree problem. Technological resources used in the study came across as powerful semiotic modes that afforded problem analysis (e.g., the interface of mobile app attending to focused features and the rich layering of realities afforded by the AR content). This study adds to our understanding of mobile technologies used as a learning tool within the context of informal science education when analyzed as a collective, learners take control of their own learning paths as they coordinated interaction between sensory modes (e.g., verbal, gaze, touch, spatial) and mobile technologies (e.g., iPads, AR).

References

Baldry, A. & Thibault, P.J. (2006). *Multimodal Transcription and Text Analysis.* London: Equinox

Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A., (Eds.) (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: National Academies Press.

Derry, S. J., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J., Sherin, M. G., & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences*, 19(1), 3-53.

Heath, C., Knoblauch, H., & Luff, P. (2000). Technology and social interaction: the emergence of 'workplace studies'. *The British Journal of Sociology*, 51 (2), 299-320.

Heath, C. & Hindmarsh, J. (2002). Analyzing Interaction: video, ethnography and situated conduct. In Tim May (Ed.), *Qualitative Research in Action*. London: Sage.

Hsi, S. (2003). A study of user experiences mediated by nomadic web content in a museum. *Journal of Computer Assisted Learning*, *19*(3), 308-319.

Hodge, R. & Kress, G. (1988). *Social Semiotics*. Ithaca, NY: Cornell UP

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39-103.

Kahr-Højland, A. (2011). Hands on, mobiles on: The use of a digital narrative as a scaffolding remedy in a classical science centre. *MedieKultur. Journal of media and communication research*, *50*, 66-83.

Kearney, M., Schuck, S., Burden, K., & Aubusson, P. (2012). Viewing mobile learning from a pedagogical perspective. *Research in Learning Technology*, *20*(14406), 1-17 . DOI: 10.3402/rlt.v20i0.14406

Kellah, E. (2000). Preparing Problem Solvers for the 21st Century through Problem-Based Learning. *College Teaching*, 48(2), 55-60

Kress, G., (2005). Gains and Losses: New Forms of Texts, Knowledge, and Learning. *Computers and Composition*, 22, 5–22.

Kress, G., Charalampos, T., Jewitt, C., & Ogborn, J. (2001). *Multimodal Teaching and Learning: the rhetorics of the science classroom*. Continuum: New York

Lemke, J. (2002). Multimedia Genres for Scientific Education and Science Literacy. In M. J. Schleppegrell & C. Colombi (Eds.), *Developing Advanced Literacy in First and Second Languages: Meaning with Power* (pp. 21-44). Mahwah, NJ: Erlbaum.

Looi, C.-K., Seow, P., Zhang, B., So, H.-J., Chen, W., & Wong, L.-H. (2010). Leveraging mobile technology for sustainable seamless learning: a research agenda. *British Journal of Educational Technology*, *41*(2), 154–169. doi:10.1111/j.1467-8535.2008.00912.x

MacArthur Foundation. (2006). *Confronting the Challenges of Participatory Culture: Media Education of the 21st Century*. Retrieved from http://www.nwp.org/cs/public/download/nwp_file/10932/

Mann, S. & Reimann,P. (2007). *Mobile Technology as a Mediating Tool for Learning in the Convergences from Technology, Collaboration and Curriculum Perspectives*. Making the connections. Conference Proceedings mLearn Melbourne 2007. Retrieved from http://www.mlearn.org/mlearn2007/files/mLearn_2007_Conference_Proceedings.pdf

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded source book* (2nd ed.). Thousand Oaks, CA: Sage.

O'Halloran, K.L. (2009). *Multimodal Analysis and Digital Technology*. Interdisciplinary Perspectives on Multimodality: Theory and Practice. Proceedings of the Third International Conference on Multimodality, Palladino, Campobasso, 2009. Retrieved from http://multimodal-analysis-lab.org/_docs/pubs02-Multimodal_Analysis_and_Tech.pdf

Pea, R. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions* (pp. 47-87). New York: Cambridge University Press.

Pea, R. (2006). Video as data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices. In Weiss et al. (Eds). *The International Handbook of Virtual Learning Environments*, 1321-1393. Springer: the Netherlands

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage

Sung, Y.-T., Hou, H.-T., Liu, C.-K., & Chang, K.-E. (2010). Mobile guide system using problem-solving strategy for museum learning: a sequential learning behavioral pattern analysis. *Journal of Computer Assisted Learning*, *26*(2), 106-115. doi:10.1111/j.1365-2729.2010.00345.x

Thomas, D & Brown, J. S. (2010). Learning in/for a World of Constant Flux: Homo Sapiens, Homo Faber & Homo Ludens revisited. In Luc E. Weber James J. Duderstadt (Eds.), *University Research for Innovation* (pp. 321-336). France: Economica.

Thomas, D. & Brown, J. S. (2011). *A New Culture of Learning: Cultivating the Imagination for a World of Constant Change*. USA: Create Space Publishing

White, T. & Pea. R. (2011). Distributed by design: On the promises and pitfalls of collaborative learning with multiple representations. *The Journal of the Learning Sciences*, 20 (3), 489-547

Yew, E.J. & Schmidt, H.G. (2012). What students learn in problem-based learning: a process analysis. *Instructional Science*, 40 (2), 371-395.

Zimmerman, H. T., Land, S. M., McClain, L.R., Mohney, M. R., Choi, G-W., & Salman, F. H. (2013). Tree Investigators: Supporting families and youth to coordinate observations with scientific knowledge. *International Journal of Science Education.* Advance online publication. DOI: 10.1080/21548455.2013.832437. pp. 1-24.

Acknowledgments

The authors express their appreciation to our partners: the Arboretum at Penn State, the Penn State Center for Online Innovation in Learning, and the Penn State Education Technology Services unit Teaching and Learning with Technology. We acknowledge the contributions of our Augmented and Mobile Learning Research Group (http://sites.psu.edu/augmentedlearning/): colleagues Lucy R. McClain, Michael R. Mohney, GiWoong Choi, and Brian J. Seely.

Promoting Student Learning through Automated Formative Guidance on Chemistry Drawings

Anna N. Rafferty, Computer Science Division, University of California, Berkeley, CA, USA
rafferty@cs.berkeley.edu
Libby Gerard, Graduate School of Education, University of California, Berkeley, CA, USA
libby.gerard@gmail.com
Kevin McElhaney, SRI Education, Menlo Park, CA, USA, kevin.mcelhaney@sri.com
Marcia C. Linn, Graduate School of Education, University of California, Berkeley, CA, USA
mclinn@berkeley.edu

Abstract: We investigated the effect of automated guidance on student-generated chemistry drawings in computer-based learning activities. Expert teachers provide guidance on generative tasks such as drawings or essays that encourages students to refine their understanding, often by gathering more evidence. We developed algorithms to score student drawings and designed guidance for each score level. The guidance was intended to promote coherent understanding. We compared computer-generated guidance to teacher guidance in two studies, conducted with over 300 students in secondary classrooms. The studies suggest that automated guidance is as effective as teacher guidance for improving student understanding. Teachers appreciated the assessment of class progress provided by the automated guidance. They reported that it took them several hours to grade their five classes of 30 to 40 students. Thus, automated guidance can reduce the time teachers spend evaluating student work, creating more time for planning lessons, facilitating inquiry, or guiding individual students.

Computer-assisted education has the potential to deliver timely guidance adapted to each student's individual ideas. Human tutors provide adaptive guidance by prompting learners to reconsider and revise their ideas, verify and elaborate on the correctness of ideas and consider ways to improve understanding (Merrill, Reiser, Ranney, & Trafton, 1992). Providing adaptive guidance is an important goal in designing computer tutors (e.g., Anderson, Boyle, Farrell, & Reiser, 1987; Anderson, Corbett, Koedinger, & Pelletier, 1995). While the majority of computer tutors provide formative guidance (Koedinger, Anderson, Hadley, & Mark, 1997; Slotta & Linn, 2009), it is often limited to student work on selection tasks (e.g. multiple-choice) or algebraic expressions. This paper explores the effect of automated, adaptive guidance on a generation task where students make drawings of chemical reactions as part of a web-based inquiry science unit.

Compared to the limited number of correct responses to selection tasks, generation tasks can adapt guidance to a wide range of student responses. Selection tasks often encourage students to recall facts rather than distinguish among ideas, and rarely provide opportunities for deep student inquiry (Shepard, 2000). Generative tasks, in contrast, elicit students' range of ideas and can encourage them to use evidence to sort out their ideas in order to create a coherent explanation. Mintzes, Wandersee, and Novak (2005) note that generative assessments can provide a fuller picture of students' conceptual understanding and drive students towards "making meaning'" rather than memorizing facts. Generative tasks can be difficult to evaluate due to the variety of responses and innumerable ways for students to express the correct answer. Previous research has found that due to the demands required in evaluating generative assessments, it is often challenging for teachers to provide detailed guidance to all students (Black & William, 1998; Ruiz-Primo & Furtak, 2007).

In this paper, we explore the effect of automated, adaptive guidance for student-generated drawings of chemical reactions as they interact with a web-based inquiry science unit. Drawings provide students with a way to express their understanding of atoms and molecules in chemical reactions (Chang, Quintana, & Krajcik, 2010). We support automated analysis of drawings by asking students to use a computer-based interface that features virtual atom stamps, rather than enabling open-ended drawings. This limits the degree to which student drawings can vary while still allowing for expression of multiple conceptual views. We designed an algorithm that diagnoses both normative and alternative chemistry conceptions in students' drawings, allowing us to align formative guidance with these conceptions. Our guidance design addresses the gap between students' ideas and the learning goal by prompting students to build on productive ideas they have, develop criteria for evaluating their own understanding, and revisit key concepts.

We explore the effectiveness of automated guidance through two classroom studies. The first investigates whether the automated guidance is as effective for promoting student learning as teacher-generated guidance. Comparing these two types of guidance allows us to identify key characteristics of effective teacher guidance, potentially informing future revisions to the automated guidance, and provides data for determining whether it is feasible to remove some of the demands on teachers by providing students with automated

 © ISLS

guidance. In this study, students receive automated guidance immediately, while they must wait until the next day to receive teacher-generated guidance, mirroring a typical classroom. The second study explores the role of immediacy by comparing delayed to immediate automated guidance. Together, these two studies address optimal ways to provide automated guidance on student drawings within the classroom and investigate how this guidance compares to that provided by teachers.

Designing Guidance for Inquiry

Guidance for generative assessments can help students improve their understanding and recognize gaps or inconsistencies in their ideas (Hattie and Timperley, 2007). It can promote learning by encouraging students to reconsider their ideas, building on their prior knowledge (Azevedo & Bernard, 1995). Generative assessments can be used to help teachers recognize students' level of understanding and adapt instruction. Ruiz-Primo and Furtak (2007) found that teachers' guidance on generative activities in an inquiry investigation was related to their students' science learning, suggesting that this monitoring can indeed help teachers boost student learning.

Providing guidance on generative assessments during instruction is difficult. Teachers often lack time to provide detailed guidance for all students on these assessments (Black & William, 1998; Ruiz-Primo & Furtak, 2007). Further, the specificity that effective guidance should exhibit remains unclear. For instance, generic guidance may prompt students to self-explain and generate their own insights (Butler & Winne, 1995; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Schmidt & Bjork, 1992) but may also allow non-normative ideas to persist (Koedinger & Aleven, 2007). Students with different levels of prior knowledge benefit from different levels of information specificity (Shute, 2008) and scaffolding (Razzaq & Heffernan, 2009). Due to the challenges of assessing and guiding student work on generative activities, selection tasks with one correct answer, such as multiple-choice questions, are the norm in science instruction. These activities are limited in their ability to capture the complexity of students' ideas.

New technologies offer promise for implementing guidance strategies to support inquiry learning. In AutoTutor (Graesser et al., 2004), for instance, a computer avatar leads a tutorial dialogue as a student solves a challenging physics word problem. The avatar prompts for more information, elicits questions, identifies and corrects "bad answers," answers the learner's questions, and summarizes responses. Across different domains and comparison groups the average learning gain is approximately one letter grade. Automated guidance has also proved effective in helping students to develop concept maps in different middle school science domains and use evidence to strengthen the links among their ideas (Segedy, Kinnebrew, & Biswas, 2013). More recently, researchers have employed machine learning techniques to automatically recognize effective inquiry practices (Sao Pedro, Baker, Gobert, Montalvo, & Nakama, 2013).

Our work adds to this body of literature on automated guidance in inquiry science by examining the effectiveness of automated guidance for drawing tasks in which students pictorially represent scientific ideas. We compare automated guidance with teacher-generated guidance, allowing us to explore what types of guidance teachers provide to students. For the automated guidance, we designed guidance messages based on knowledge integration principles. Knowledge integration is based on constructivist ideas that focus on building on students' prior knowledge and helping them to connect new concepts with this knowledge, even if some of this prior knowledge is non-normative (e.g., Smith III, Disessa, & Roschelle, 1994). Knowledge integration guidance can assist students by prompting them to compare and contrast their views with evidence or scientific theories, or to add new ideas missing from their initial conception (Linn & Eylon, 2006). When guidance directly builds on students' own ideas, as articulated in their initial response to the activity, it can help students develop criteria for distinguishing between normative and non-normative ideas and push students to integrate ideas rather than holding separate, conflicting conceptions (Linn & Eylon, 2011).

Curriculum and Drawing Activity

We focus our investigation of automated guidance on students' drawings of chemical reactions. The drawing tasks are part of an inquiry unit in the Web-Based Inquiry Science Environment (WISE), entitled *Chemical Reactions: How Can We Help Slow Climate Change?* (Chiu & Linn, 2011). *Climate Change* addresses difficulties students have with understanding chemical reactions as composed of discrete particles (Ben-Zvi, Eylon, & Silberstein, 1987) by highlighting the role of conservation of mass, ratios, and excess reactants in combustion reactions.

Past work has shown that learning multiple representations of chemical reactions and providing students with ways of visualizing the particles in reactions can help to strengthen understanding (Harrison & Treagust, 2000; Schank & Kozma, 2002). The drawing tasks ask students to draw the arrangement of atoms before and after a chemical reaction [Figure 1(a)]. One of the tasks focuses on the combustion of methane and the other ethane. The WISE Draw screen provides students with "stamps" for each atom; for instance, the methane reaction includes stamps for oxygen, carbon, and hydrogen. Students must choose how many of each atom to add to their drawing and arrange the atoms to reflect how they are grouped into molecules. The draw interface allows students to create multiple *frames*, one to show the atoms before the chemical reaction

(reactants) and one to show the atoms after the chemical reaction (products) [Figure 1(b)]. After creating a new frame, students rearrange the atoms to show the products of the reaction. This activity presents chemical reactions as the rearrangement, rather than the creation or destruction, of matter. Students may still add or delete atoms in a way that reflects their conceptual misunderstandings about conservation of mass. The drawings enable students to articulate their ideas about chemical reactions, while constraining the representation to enable automatic evaluation. The drawings provide an opportunity to work with a different model of chemical reactions than the typical equation-based format, and students frequently demonstrate non-normative ideas in the activity. Our goal is to provide conceptual guidance targeting non-normative or missing ideas in the students' drawings.

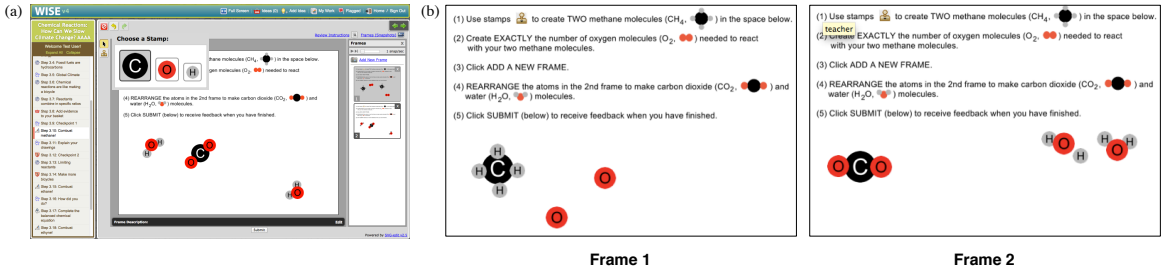


Figure 1. Drawing and feedback interface. (a) The WISE Draw interface. Students use "stamps," the black, red, and gray circles, to represent atoms and show the molecules before and after a chemical reaction. (b) The frames of a student drawing. The scorer recognizes that CH_4, CO_2, and H_2O are allowed molecules, but that the student has not conserved mass and has placed two separate oxygen atoms rather than an oxygen molecule.

Evaluating Student Drawings

To evaluate student drawings, we created an algorithm that processes each drawing and assigns a score. Based on examination of 98 drawings from past students, half methane and half ethane, we identified common student ideas and grouped these ideas into conceptual categories, shown in Figure 2(a). Each category includes a different conceptual feature, such as conserving mass from the beginning to the end of the reaction or correctly representing the reactants. The concepts are organized into a hierarchy from more basic to more complex. We evaluated the accuracy of the algorithm on the development set of 98 drawings as well as on a test set of 200 additional drawings from past students. Both sets of drawings were scored by a trained human scorer, and the test set was not examined until after the algorithm was developed. For the development set, the algorithm's score matched the human's score for 96.9% of the drawings; for the test set, the scores matched in 91.5% of the drawings. This compares favorably with other systems for scoring student answer to generative assessments; for instance, the C-Rater system scores short answer responses, including responses to science prompts, and matched human scores about 84% of the time in two separate evaluations (Leacock & Chodorow, 2003). Note that while development focused on the methane and ethane tasks, the system can score other chemical reactions drawings. Information about the drawing task is provided as an XML input file specifying the correct molecules in each frame, allowing the scoring algorithm to be agnostic to the specific task being scored.

(a)

Criteria	0	1	2	3	4	5
Two frames		✓	✓	✓	✓	✓
Conserves atoms			✓	✓	✓	✓
Reactants correct				✓	✓	✓
Products correct					✓	✓
Groupings clear						✓
Rate in dev. set	11%	19%	16%	5%	3%	45%

(b)

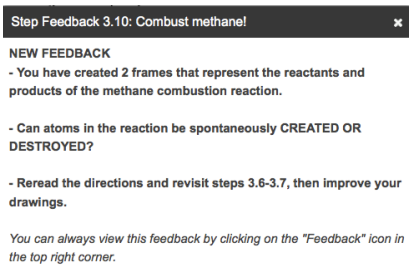


Figure 2. Drawing rubric and automated guidance window. (a) The automated scoring rubric. Drawings receive the highest score for which they exhibit all checked criteria. (b) An automated guidance window. Students are given personalized guidance, designed to promote knowledge integration, based on the automated scoring rubric. This guidance message corresponds to a score of 1, where the drawing has not conserved mass.

Creating Guidance from Scores

Given the scorer's ability to accurately evaluate student drawings, we can provide guidance based on the conceptual understanding that the student has. For each of the six possible scores, we designed a textual message to help students revise their drawing. The textual guidance was designed to promote knowledge integration by recognizing students' normative ideas and helping them to refine and revise their non-normative ideas (Linn & Eylon, 2011). Drawings that were scored as having some conceptual error (scores 0-4) all

received textual feedback of a similar format. First, a correct feature of the drawing was recognized, anchoring the guidance with students' prior knowledge. For example, a student whose drawing received a score of 2 would be acknowledged for conserving mass, since this is the most complex conceptual feature exhibited by the drawing. The textual feedback then posed a question targeting the student's conceptual difficulty, such as identifying what molecules should be present in the reactant frame; this elicits student ideas about the topic of difficulty. Finally, the feedback directed students to a relevant step earlier in the unit, and encouraged them to review the material in that step and then revise their drawing. This promotes adding new ideas and distinguishing normative and non-normative ideas. Figure 2(b) shows the guidance for a score of 1.

Study 1: Comparing Student Learning with Teacher vs. Automated Guidance

We compare automated knowledge integration guidance to teacher-designed guidance on students' chemistry drawings. If computer-selected knowledge integration guidance could help students improve, the computer could save the teacher valuable time, which could be spent planning instruction and working with individual students. Computer-assigned guidance differs from teacher-generated guidance in several ways. The teacher-generated guidance can be customized based on the teacher's knowledge of individual students. For example, a teacher might respond to a conservation of mass error differently if the drawing was made by a student who typically struggles in science versus a student who typically excels. The teacher is also likely to be able to differentiate conceptual misunderstandings from errors due to use of the interface. The timing of the guidance also differs. Automated guidance is provided to students immediately, allowing them to revise their understanding before moving on in the unit. Previous studies suggest mixed results for immediate and delayed guidance (Shute, 2008). Given that immediacy is a unique affordance of computer-assigned guidance, we wanted to test its value. Overall, we hypothesize that the automated guidance will be as effective as teacher-generated guidance for promoting student learning given that the automated guidance targeted common conceptual errors and was designed to promote cohesive integration of ideas.

Methods

Participants

Eighth grade physical science students (N=263 completed both pre- and post-test, N[groups]=129; ages typically 13-14) from 10 classes in a public middle school participated in the study. Classes were taught by one of two teachers; each teacher taught five of the ten classes. Teachers were selected who have over three years experience teaching the WISE *Climate Change* unit and writing guidance for student work on the drawing tasks.

Study Design and Administration of Feedback.

Students were randomly assigned by class period to receive either automated or teacher-generated guidance. Three class periods from each teacher were assigned automated guidance (AG), and the other two periods teacher guidance (TG). Students completed the *Climate Change* unit in the classroom as part of the curriculum.

Students in both conditions completed the same pre-test prior to *Climate Change.* The same items were administered as a post-test after *Climate Change.* Pre- and post-tests were completed individually.

Students worked through *Climate Change* in groups of one to three students. All students experienced the same activities in *Climate Change* except for the draw steps. The two draw steps occurred in a part of the unit focusing on combustion reactions and their contributions to climate change. In the draw steps, all students received the same instructions about the use of the WISE Draw interface and the chemical reaction to depict. Students in the automated condition were told to click the "Submit" button when they wished to receive feedback. When students clicked this button, they were warned that they only had two chances to receive feedback and to confirm that they wanted to proceed. After confirming, a pop-up box with the textual feedback appeared. Students could close the feedback or re-open it to view their existing feedback at any time. If students clicked the "Submit" button more than twice, they were told that they had used all of their opportunities to receive feedback, but that they could continue to revise their drawing if they wished.

In the teacher guidance condition, the researchers met with the teachers before the start of the unit to review previous student drawings on this item and discuss a possible scoring rubric and guidance approaches. After the unit started, the teachers reviewed student work at night after all students had completed the drawing steps, and wrote guidance. Students received the guidance at the start of class the next day when they logged into the WISE unit. After logging in, a pop-up told students they had received guidance from their teacher on the draw task and provided them a link to jump immediately to that step in the unit. The guidance interface on the draw steps was identical to the interface in the automated condition.

The two conditions differed in how many rounds of guidance were provided: students in the automated guidance could receive up to two guidance messages, while students in the teacher-generated condition could receive only one. Automated guidance can be provided multiple times without significantly lengthening the unit; providing multiple rounds of teacher guidance, given that teachers need to review student work after class, is

infeasible given the number of days available for students to work on *Climate Change* and the location of the drawings in the activity. While this distinction means that we cannot draw conclusions about whether teacher-generated guidance would be more effective if multiple rounds were provided, we believe this design better reflects the way that guidance would actually be available to students and thus can provide evidence for the best way to incorporate guidance in classroom activities.

Data Sources and Scoring of Knowledge Assessments

Measures of learning included students' initial and revised drawing for each of the methane and ethane drawing tasks, and an item from the pre- and post-tests. The pre-post test item called for students to transfer ideas learned in the drawing tasks, critiquing a molecular drawing made by a hypothetical peer for a new chemical reaction formula. A knowledge integration rubric was used to score student responses on the drawings and the item on the pre- and post-tests. The five-point scoring rubric evaluated student response in terms of integrating ideas about conservation of mass, ratios and excess in chemical reactions. Unlike the automated scoring system, this rubric was not hierarchical. We also conducted teacher and student interviews during enactment, documented the teacher-generated guidance, and collected rich classroom observation notes.

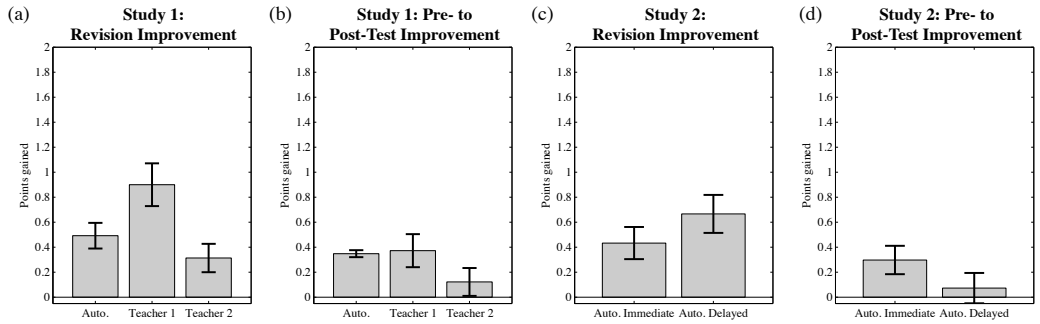


Figure 3. Improvements based on guidance type. Error bars show one standard error. (a) Revision improvement in Study 1. (b) Pre- to post-test improvement in Study 1. (c) Revision improvement in Study 2. (d) Pre- to post-test improvement in Study 2.

Results

Overall, the automated guidance was as helpful for student learning as the teacher-generated guidance. Students made modest improvements from their initial drawings to their final drawings, increasing their scores by an average of 0.56 points (Cohen's d=.36). Students in both the AG and TG conditions showed similar amounts of improvement (d=.33 and d=.39, respectively). A repeated measures analysis of variance (ANOVA) with factors for condition and initial versus final drawing showed significant improvement after revising the drawing ($F(1,481)=50.36$, $p<.001$) but no significant interaction between condition and amount of improvement ($F(1,481) = 0.85$, $p > .3$).

Students also showed improvement from pre- to post-test on the transfer item in which they critiqued the drawing of a hypothetical peer. Students in both conditions showed similar improvements in performance on this item (AG: $t(154)=4.63$, $p < .001$, d=.42; TG: $t(107)=2.93$, $p < .01$, d=.36). A repeated measures ANOVA showed that there was no significant interaction between guidance type and amount of improvement.

Analysis of the teacher-generated guidance demonstrated substantial differences in the quality of guidance given by the two teachers. One teacher wrote substantially more detailed comments that focused on both chemistry concepts and features of the drawings, whereas the other teacher wrote more terse comments; both teachers mentioned that providing guidance to students took several hours in the evening.

Students of the teacher who wrote more conceptual comments made significantly greater improvements on their drawings. An analysis of variance on the amount of improvement in drawing scores from initial drawing to final revision, with a factor for guidance type (automated, Teacher 1, or Teacher 2) and a random factor for student group showed that guidance type had a significant effect on amount of improvement ($F(2,481) = 4.04$, $p < .025$). As Figure 3(a) shows, students who received more conceptual guidance (Teacher 1) improved more than students in other conditions did, and students who received automated guidance improved more than students who received terse guidance (Teacher 2). While this interaction was not significant for pre- to post-test improvement, the same trend held: students who received guidance from Teacher 1 improved an average of 0.37 points (d=.48), students in the automated condition improved 0.35 points (d=.42), and students who received guidance from Teacher 2 improved 0.12 points (d=.20; see Figure 3b).

The teacher who wrote more conceptual comments used a relatively small number of comments for all students, customizing these comments slightly on a case-by-case base. Each comment focused on a particular conceptual issue. For example, one comment was *"You have only made one frame to represent the products and reactants. Your first frame should be for the reactants. A second frame should be made for the products.*

Follow the directions on the top of the page." In contrast, the second teacher gave short comments that were often solely procedural, such as directing students to read the directions; this teacher commented that he had little time to review the drawings. Conceptual comments from this teacher tended to state a concept in isolation, such as *Conservation of mass?"* These comments may have been too terse to help students integrate these concepts into their revised drawings.

Student interviews point to both challenges and benefits of automated guidance in terms of helping students to monitor their learning. The automated guidance may provide students an alternative to relying on the teacher for answers as several students noted , "*[Teachers]'ll be more specific. They'll show you. They'll point out what's wrong*." while the automated guidance encouraged students to sort through their ideas on their own. While a number of students wished that the automated guidance would "*tell the person exactly what they did wrong*," one pair noted the benefits of less specific guidance: "*If you don't get a problem, the teacher may give the answer away…they will be like 'No, it's like this,' and they will do it for you. But you need to learn for yourself.*" On the other hand, the automated guidance was unable to provide the extended dialogue that teachers may facilitate to ensure students grapple with the concepts *and* reach an understanding. A combination of teacher and automated guidance may provide the best solution for promoting student learning.

Study 2: Timing of Guidance

In Study 1, the automated guidance was as effective as teacher-generated guidance in helping students revise their drawings and improve their post-test performance. Students received these two types of guidance at different timing intervals to reflect typical use – immediately from the computer and start of class the next day from the teacher. This variation in timing may have contributed to the effects of the automated and teacher-generated guidance on learning. Study 2 examines the question of whether the benefit of automated guidance is tied to its immediate timing. Previous studies have found mixed results concerning the effectiveness of immediate versus delayed guidance (for a review of the literature, see Shute, 2008), with evidence that particular learner and task characteristics may influence which method is more effective. Immediate guidance is often more effective, especially for struggling students and on more challenging tasks (Shute, 2008). For our tasks, we hypothesized that immediate guidance would be more engaging and motivating to students, and would allow them to improve their understanding prior to moving on to related tasks in the unit.

Methods

Participants

Ninth grade basic chemistry students (N=88 completed both pre- and post-tests, N[groups]=57; ages typically 14-15) from four classes in a public middle school participated. The same teacher taught all students.

Study Design and Administration of Feedback

Students were assigned to the immediate or delayed guidance conditions on a full-class basis. All students completed pre- and post-tests individually, and completed the WISE unit in groups of one to three students.

The immediate guidance condition was identical to the automated condition in Study 1. We provided guidance to students in the delayed condition the evening after they completed their initial drawings. When students returned the following day, they were informed that they had new guidance and viewed the guidance. In both cases, the comments were based on the score of their drawing, and the texts in the two conditions were identical. Students in the immediate guidance condition could submit their drawing up to two times; due to time constraints, students in the delayed condition received only a single round of automated guidance.

Data Sources and Scoring of Knowledge Assessments

We evaluated student drawings (initial and final, after revisions) and the pre- and post-test item using the same knowledge integration rubrics as in Study 1.

Results

Overall, the outcome measures showed similar learning regardless of guidance timing. Students in the immediate condition improved their drawings by an average of 0.43 points (d=.36) compared an average of 0.67 points improvement for students in the delayed condition (d=.49) [Figure 3(c)]. A repeated-measures ANOVA including factors for revision (initial versus final) and guidance condition, as well as an interaction between these two factors, showed a main effect of revision ($F(1,227)$=25.5, $p < .001$), but no significant effect of condition or of the interaction. On the post-test item, students showed small, reliable improvements from their pre-test scores, with an average improvement of 0.19 points (d=.27). A repeated measures ANOVA with factors for pre- versus post-test and feedback condition showed that both main effects were significant (pre- versus post test: $F(1,86) = 4.58$, $p < .05$; condition: $F(1,86) = 4.12$, $p < .05$). Closer examination revealed relatively little improvement for students in the delayed condition (d=.10) compared to students in the immediate condition

(d=.43) [Figure 3(d)]; however, an analysis of covariance did not show a significant difference in gains when pre-test scores were included as a covariate ($p > 0.3$).

These results suggest that the benefit of automated guidance is not simply due to its immediacy, although further study is needed to determine if the two guidance types lead to differences in retention. While this study illustrates that automated guidance can also be effective when delayed, facilitating a variety of classroom implementations, immediate guidance is likely to be the most common approach. Immediacy is a unique affordance of computer-assigned guidance relative to teacher guidance. It can advance students during class enabling teachers to work with those who continue to struggle. Further, immediate guidance can be more easily integrated into activities and provides an intuitive appeal by helping students develop understanding of challenging material before moving on.

Discussion

Formative guidance can help students to distinguish between normative and non-normative ideas in generation tasks by building on their current knowledge. Our automated system accurately evaluated student responses and provided guidance on their chemistry drawings. The knowledge integration guidance focused students on comparing alternatives and integrating normative concepts to strengthen their own understanding. Of course, not all improvements on the pre- and post-test item were solely the result of the guidance. But, students' revisions suggest that they used the hints provided to revise their understanding. We compared providing automated guidance on demand or at a delay and found that these approaches were equally effective in helping students revise their understanding of chemical reactions. Overall, these studies show that student generated drawings with a wide range of responses are amenable to automatic evaluation. They also show that a small set of knowledge integration guidance options can encompass a wide range of student responses. Compared to guidance on selection items where students often just change the response to the correct one, guidance on these generation items motivated students to review their own work and identify revisions.

Overall, the automated knowledge integration guidance was as effective as teacher-generated guidance for promoting learning. The effect of automated guidance in the two teachers' classrooms suggests different roles that automated guidance may play to augment inquiry learning. In Study 1, we found that teachers differed in the type of guidance that they provided and that this lead to differences in student outcomes. For one teacher, the automated guidance was significantly more effective than their own written guidance, likely because this teacher had little time to write individualized conceptual guidance for each student drawing. In this classroom, the automated guidance could serve as a starting point for discussions between teacher and student by identifying conceptual issues that need refinement. For the teacher who had time at night to review students' work, automated guidance left him with less knowledge about his students' ideas, making it more challenging to plan instruction for the next day. In this classroom, automated guidance may work best as a tool to help the teacher plan instruction. The teacher could review the automated guidance and scores at night, and sort student work by auto-score to see categories of student ideas. This would enable the teacher to customize instruction and save valuable time reviewing responses one by one. As automated guidance becomes a part of computer-based curriculum, it is essential to work with teachers to customize use of the information.

Our work is consistent with a growing body of research suggesting that effective formative guidance can be provided for generative items within educational technologies. We are extending this work by exploring varied types of formative guidance. Textual guidance is typically provided by teachers because designing customized activities for each student is prohibitively time consuming. However, adaptive activities based on students' drawings could prove to be more engaging and lead to richer student insights than simple text messages. The automated scoring system provides a general tool for testing different types of adaptive guidance based on students' drawings. Our investigations provide support for the use of automated guidance in the classroom. These results also point to the potential of automated guidance and scoring for generative items in cases where teacher-generated guidance is infeasible, such as online courses with thousands of students. To design our scoring and guidance system, we relied on analysis of previous student work as well as educational principles, including the knowledge integration framework. This design process provides an example of how automated guidance can be created for new items and revised as additional student work is collected. Overall, our work demonstrates the effectiveness of automated guidance for student-created drawings and provides a new tool for exploring how best to deploy this guidance to help students and teachers.

References

Anderson, J., Boyle, C., Farrell, R., & Reiser, B. (1987). Cognitive principles in the design of computer tutors. *Modelling Cognition*, 93–133.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167–207.

Azevedo, Roger, and Robert M. Bernard. "A meta-analysis of the effects of feedback in computer-based instruction." *Journal of Educational Computing Research* 13.2 (1995): 111-127.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students' visualization of a chemical reaction. *Education in Chemistry*, *24*(4), 117-120.
Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In *Metacognition: Knowing about knowing* (p. 185-205). Cambridge, MA: The MIT Press.
Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74.
Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281.
Chang, H. Y., Quintana, C., & Krajcik, J. S. (2010). The impact of designing and evaluating molecular animations on how well middle school students understand the particulate nature of matter. *Science Education*, *94*(1), 73-94.
Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145-182.
Chiu, J., & Linn, M. (2011). Knowledge integration and wise engineering. *Journal of Pre-College Engineering Education Research (J-PEER), 1*(1), 1-14.
Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 180-192.
Harrison, A. G., & Treagust, D. F. (2000). Learning about atoms, molecules, and chemical bonds: A case study of multiple-model use in grade 11 chemistry. *Science Education, 84*(3), 352–381.
Hattie, J., & Timperley, H. (2007). The power of feedback. Review of educational research, 77(1), 81–112.
Koedinger, K., Anderson, J., Hadley, W., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*(1), 30–43.
Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239–264.
Leacock, C., & Chodorow, M. (2003). C-rater: Automated scoring of short-answer questions. *Computers and the Humanities*, *37*(4), 389-405.
Linn, M. C., & Eylon, B. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration.* Routledge.
Linn, M. C., & Eylon, B. S. (2006). Science education: Integrating views of learning and instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology, 2nd edition* (p. 511-544). Mahwah, NJ: Lawrence Erlbaum Associates.
Merrill, D., Reiser, B., Ranney, M., & Trafton, J. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences, 2*(3), 277–305.
Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2005). *Assessing science understanding: A human constructivist view*. Academic Press.
Razzaq, L. M., & Heffernan, N. T. (2009). To Tutor or Not to Tutor: That is the Question. In *Proceedings of Artifical Intelligence in Education,* 457-464.
Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.
Sao Pedro, M. A., Baker, R. S. de, Gobert, J. D., Montalvo, O., & Nakama, A. (2013). Leveraging machine-learned detectors of systematic inquiry behavior to estimate and predict transfer of inquiry skill. *User Modeling and User-Adapted Interaction*, 23(1), 1–39.
Schank, P., & Kozma, R. (2002). Learning chemistry through the use of a representation-based knowledge building environment. *Journal of Computers in Mathematics and Science Teaching*, *21*(3), 253–279.
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological science*, *3*(4), 207-217.
Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153-189.
Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2013). The effect of contextualized conversational feedback in a complex open-ended learning environment. *Educational Technology Research and Development, 61*(1), 71–89.
Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29(4)*, 4-14
Slotta, J., & Linn, M. (2009). *WISE science: Web-based inquiry in the classroom.* Teachers College Press.
Smith III, J. P., Disessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences, 3*(2), 115–163.
Zhang, Z. H., & Linn, M. C. (2011). Can generating representations enhance learning with dynamic visualizations?. *Journal of Research in Science Teaching*, *48*(10), 1177-1198.

Acknowledgments

This work was supported by NSF grant number DRL-1119670 to Marcia C. Linn.

Computer-Enhanced Dialogic-Reflective Discourse

Shiri Mor-Hagani and Dani Ben-Zvi The University of Haifa, Israel
Learning In a NetworKed Society, Israeli-Center of Research Excellence (LINKS I-CORE)
Email: shihag@gmail.com, dbenzvi@univ.haifa.ac.il

Abstract: Various studies on computer-supported collaborative learning have shown that attempts to implement collaborative learning in learning communities frequently encounter serious challenges. The current paper, following the philosophical approaches of Martin Buber and Nel Noddings, suggests a way to enhance students' positive appreciation of collaborative learning as well as the extent and quality of their engagement in it. To this end, we propose a preliminary conceptual framework, called Computer-Enhanced Dialogic-Reflective Discourse, which explains the dimensions of discourse in which students co-explore the challenges arising from their collaborative learning experiences. We illustrate and discuss the supposed potential of our proposed framework.

Key Words: Computer-Enhanced Dialogic-Reflective Discourse (CEDRD), Computer-Supported Collaborative Learning (CSCL), dialogue, learning community, reflection.

> "The learning community gave me a lot. From being an individual learner, I learned to appreciate the importance of the community members who shared the learning process, to appreciate my ability to learn enormously from community members, and to grow with their support." (Gali, member of the CATELT learning community)

Introduction

Previous studies of computer-supported collaborative learning (CSCL) environments have shown that efforts to implement collaborative learning in educational settings frequently encounter difficulties, such as lack of students' motivation to actively participate in online interactions, superficial or spurious collaboration (Salmon, 2000; Wegerif, 2007) and unequal contributions by students throughout the learning processes (Kerr & Bruun, 1983). Research suggests that collaborative learning should be designed more carefully and be mediated to a greater extent by teachers (Brown & Campione, 1994; Dillenbourg, 2002; Guzdial et al., 1997). In recent years, research has included a focus on socio-emotional aspects of CSCL (Bielaczyc, 2006), for example the significance of the relationships among students, emotions that accompany and affect collaborative learning and learning communities (Baker et al., 2013; Dillenbourg et al., 2009; Hod & Ben-zvi, 2013), interpersonal conflicts (Baker et al., 2013; Ben Zvi, 2007), and students experiencing a sense of alienation in online courses (Shner, 2012).

This proposal is situated within the discussion on socio-emotional aspects of CSCL. More concretely, it seeks to demonstrate how a conceptual framework, based on principles adapted from Buber's Dialogic philosophy (Buber, 2007) and from the Ethic of Care perspective (Noddings, 1995; 2012) can contribute to this discussion. In this paper we therefore introduce CEDRD, a preliminary conceptual framework of Computer-Enhanced Dialogic-Reflective Discourse. The CEDRD framework is based on a study of four rounds of a semester-long graduate course entitled Challenges and Approaches to Technology-Enhanced Teaching and Learning (CATELT, Hod & Ben-zvi, 2013) that was designed as a blended, computer-supported learning community.

We first introduce the theoretical background of the CEDRD's two core dimensions: the dialogic and the reflective. Then, we elaborate on the way they are manifest in the framework's third dimension: the design of the learning environment. Through examples, we illustrate and discuss the significance of the CEDRD framework.

Dialogue and Learning

Researchers have proposed various links between the concept of dialogue and CSCL, suggesting different understandings of relevance and applications (e.g., Burbules, 1993; Koschman, 1999; Wegerif, 2007). We believe that several principles of Buber's dialogic philosophy and Noddings' concept of moral education, may hold special importance to CSCL's research and practice in allowing us to not only discuss the significance of dialogic social relations in these settings, but also to discuss their necessary composition and nature. After a brief introduction of these perspectives, we will discuss their relevance.

Buber (2007) argues that human self-realization as an "I" emerges and develops through and within the dialogic encounter with the other (Bergman, 2007; Buber, 2007). Dialogue happens between equal participants, who perceive each other as an end in itself, mutually engaged in an open and direct relationship, which Buber

calls the "I–Thou" relationship (in contrast to "I–it" relationships, in which the other is seen as an object to be manipulated or a means to obtain other goals). The I–Thou relationship is characterized by its participants' attentiveness and responsibility toward one another, as well as their mutual confirmation of their acceptance, affirmation and support of the other in its otherness (Buber, 2007; Kramer, 2013). Central to the possibility for a true dialogue to occur is what Buber called inclusion – meaning, making present (Kramer, 2013), or "experiencing the other side to feel an event from the side of the person one meets as well as from one's own side" (Friedman, 1956, p. 96). It is through inclusion that one can perceive and recognize the others in their full humanness, as they really are. Furthermore, it is through inclusion that one can grasp, not only the other's character and talents, but a view of the situation and of the other. This may happen only through opening up oneself to the other's otherness without forgetting or relinquishing what he himself is (Buber, 2007; Friedman, 2002).

The educational dialogue differs from other I–Thou relationships in that the teacher and the student are engaged in one-sided inclusion (Friedman, 1956). Based on the notion that this asymmetry is central to the teacher-student relationship, the student does not and cannot experience the side of the teacher. Yet it is through inclusion that the teacher knows whether she acts in a manner that is appropriate for and serves to benefit her student.

Noddings, one of the most influential scholars of the ethic of care perspective (Alpert, 2008), emphasizes the importance of creating caring relations within the classroom. Noddings distinguishes between the teacher's role as carer, and students' role as the cared-fors. This occurs when the teacher directs a receptive attentiveness (which Noddings described as containing and tolerance attention) to the needs expressed by students (Alpert, 2008; Noddings, 2007) and to the students' worlds. Then, the teacher has to respond to these needs, either by addressing them or, in cases of conflict between a student's needs and the needs of the education system in which both exist, by ensuring that caring relations between them will continue and not be harmed by the teacher's inability to address these needs (Noddings, 2012).

Noddings connects Buber's dialogic philosophy and the ethic of care perspective (Noddings, 1995; 2010; 2012). It is through and within dialogue that students can get engaged in discussion about the meaning of care, and reflect and critically analyze what may or may not be considered as caring behavior for themselves and towards others. In addition, dialogue contributes to the development of the cared-fors since it is through dialogue that the teacher may invite her students to "examine their own lives and explore the great questions human beings have always asked" (Noddings, 1995, p. 191). Furthermore, it is through dialogue that the teacher may learn about her students, their worlds and their needs, to better direct her efforts as carer, as well as getting from them the necessary feedback regarding her efforts (Alpert, 2008).

The relevance and special significance of these two perspectives stem from their emphasis on the relationships among learning community members as the heart of the educational process and its impetus, and as the key factor affecting students' engagement in this process. In the context of the proposed CEDRD framework, the importance of these perspectives comprises more concretely of: a) an acknowledgement of the central role of the dialogic communication, characterized by care, attentiveness and confirmation; b) an emphasis on participants' deep, comprehensive, and frequently reworked knowing of each other; c) an understanding of the students' position as equal in value yet different in essence from the position of the teacher; and d) an emphasis on the students' expressed appreciation of the caring relationships and their trust in the teacher.

Reflective Interactions in CSCL Communities

Reflective interaction, considered as an interaction in which partners not only propose solutions to a problem, but offer explanations, justifications, and/or assessments of these solutions, has been found to be one of the three interaction patterns that most effectively promote learning (Baker & Lund, 1997). Salmon (2002) described reflective interaction as a process of suggesting ideas, receiving feedback, and reforming initial ideas in response to the feedback. While these understandings of reflective interactions relate to theoretical knowledge, in the CEDRD framework the focus of reflection includes also the students themselves and the learning processes in which they are engaged. By this, we integrate the above definition with the understanding of reflection as individuals' directed critical examination of their own past conduct (including beliefs, thoughts, and emotions) for the purpose of drawing conclusions that may be implemented in the future (Bengtsson, 1995; Hatton & Smith, 1995).

The integration of the dialogic and reflective dimensions in the context of CSCL, was previously proposed by Wegerif (2007). Wegerif proposed a model comprising three partially overlapping types of dialogue, each encompassing the others to some extent: critical dialogue (solving problems and making judgments), caring dialogue (understanding the other), and creative dialogue (exploring solutions, raising arguments). In addition, this model emphasizes the opening, deepening and broadening of reflective and dialogic spaces between students that enhance creative learning and learning how to learn. Therefore, dialogue is viewed not just as a means for constructing knowledge but as an end unto itself.

While there are several similarities between this model and the CEDRD framework, we place primary emphasis on the significance of interpersonal relations and personal acquaintance and their leading role in the learning process. Another difference is the context in which the concept of reflectivity is used. For Wegerif, reflectivity entails an exploration of theoretical knowledge, perspectives, etc., while in the CEDRD framework, reflectivity first and foremost entails an exploration of the students themselves in different contexts.

The Conceptual Framework of CEDRD

Having presented the theoretical foundations of the CEDRD framework, we discuss its dialogic and reflective dimensions and their manifestation in a computer-enhanced learning environment. After elaborating on the essence of each separately, we illustrate the meaning of the integration of these dimensions into the framework.

The Dialogic Dimension: Engaging in an Attentive, Caring Learning Community

The dialogic dimension includes both the teacher's as well as the students' context (Table 1).

Table 1: The teacher and students' contexts of the dialogic dimension

The teacher's context	The students' context
1. A caring, attentive, and accepting attitude. 2. An educational approach that promotes ongoing acquaintance, attentiveness and care among learning community members. 3. Integrating social activities to deepen students' acquaintance and caring for each other. 4. Treating the conflicts that arise from collaborative interactions not as obstacles but as opportunities to deepen relationships and enhance involvement.	1. Tolerance, mutual trust, and caring for each other. 2. Sharing experiences, feelings, insights, and ideas with each other, orally and in writing. 3. Responding (orally and in writing) to other's reflections and comments. 4. A growing acquaintance with each other. 5. Engagement in an on-going discourse about the dilemmas and conflicts that emerge in the learning process.

The Reflective Dimension: 'Reflecting on' and 'Reflecting within' Collaborative Learning

The reflective dimension first of all implies the integration of reflective observation on how learning content is connected to students' lives, and on the individual, team, and collaborative community learning processes, as an integral part of learning. Moreover, the reflective dimension also implies the presence of several elements in the discourse (Table 2).

Table 2: Reflective elements in the community discourse.

The students' context	The community's context	The teachers' context
1. Students' oral and written reflective expressions throughout the course. 2. Students' mutual exposure to others' reflections.	1. Community reflective conversations regarding individual dilemmas, interpersonal conflicts and community challenges arising from collaboration. 2. Community exploration of the significance of experiencing collaborative learning in a community of learners.	1. Raising questions that are designed to clarify or refine students' reflections. 2. Modeling how and in what ways one's questions may assist another's reflective exploration.

The Design Dimension: Creating Spaces and Devoting Time for Ongoing Discourse

The blended learning environment combines in-class community reflective conversation (CRC) sessions with an online learning platform with the following vital design features (Table 3).

Table 3: Design features of the blended learning environment.

In-class community reflective conversations	Online learning platform
1. A space designed for community conversation (such as sitting face to face in a circle), that supports open, informal talk. 2. Time devoted for community conversation (Takes a substantial part of the in-class sessions' total time, and is repeated weekly).	1. A space to conduct discussions about the collaborative knowledge emerging in the collaborative learning process. 2. A space devoted to students' written reflective expressions. 3. The visibility of the students written reflections to all learning community members.

In-class community reflective conversations	Online learning platform
	4. Students' ability to comment on each other's reflections, as well as on each other's comments.

Engaging in Computer-Enhanced Dialogic-reflective Discourse

Over the four annual iterations of research in CATELT, we repeatedly observed the emergence and development of discourse among the learning community members about the socio-emotional aspects of collaborative learning. This discourse was mostly feasible in the students' diaries, located at the Wiki course website, used by students to record their experiences and reflective insights on the learning process, and the weekly CRC, which takes place in-class' face to face meetings. From the second iteration on, we used several principles of Buber's dialogic thought and Noddings' ethic of care perspective, both to analyze and to further enhance the learning community discourse. To make our description of the CEDRD somewhat less theoretical, we provide several examples that illustrate its nature and some key features. In addition, they provide an opportunity to highlight potential significance of CEDRD.

Adopting a New Perspective upon Oneself and about Learning Collaboratively

Three students are the participants of the first example: Ariel, Dalit and Michal (All the students' names have been changed to preserve anonymity). The three worked on a collaborative editing task that required them to jointly write a summative critique of a paper. They began during the face-to-face meeting, and continued in the Wiki course website. In the CRC that took place the following week, one student described that she had felt pushed aside in the course of the collaborative work. The dominant student in this group made no comment during the CRC, but later wrote the following passage in his personal diary:

> Ariel: After a brief exchange with Dalit and Michal, I accessed the editing window and, at my typical pace, started to delete, modify, and rephrase, effectively pushing aside everything that had appeared in the original article. Only after I returned home and looked at it again, did I start to feel a strong sense of discomfort – What right did I have? Who says that my interpretation of the article's meaning is better or more correct than what they had written? I looked at how I had been working in class and it seemed to be very disrespectful of the others' work, very aggressive and overbearing, maybe even violent… This did not match the image that I had of myself in my mind, which did not include being insensitive and overbearing!

Several days later, the following comment was posted on the joining Wiki conversation page:

> Michal: The easiest thing is to blame you for being overbearing and causing paralysis in those around you through the speed of your thinking and your work, but what about [my own] personal responsibility? I was not very happy with the answers that I was forced to give to myself.

The next day, another comment was posted, this time from the third student in the group:

> Dalit: For the most part I felt that you prompted me to keep going forward. I did not get the feeling that you were taking control of the computer. In general, I think that our joint learning process was too short. You have to get to know the team you are working with, and everyone has to get to feel comfortable, get to know the strengths and weaknesses of each other, among other things.

This example illustrates one of the main elements of the CEDRD framework: a collaborative dialogic reflection. By this, we refer to the manner in which the encounter with the observations and contemplations of others leads students to think about themselves and to adopt a new perspective on themselves. In this respect, the students' exchanges can be seen as a process of negotiation of the meaning of the events—whose goal is not necessarily to arrive at an agreed interpretation, but rather to reach an understanding of the personal implications that apply to each student separately, to their relations, and to their fundamental view of the nature of collaborative learning. This example demonstrates the potential of the CEDRD framework to enable and promote this kind of written communication. Moreover, throughout each of the course's iterations, collaborative dialogic reflection was manifested in two main forms: reflections expressed in the CRC were commented on later in the Wiki's personal diaries, and written reflections were mentioned and discussed in the following CRC. In some cases, we also observed how a discussion over one theme continued from CRC to the Wiki, back to the

CRC, and so on. This dynamic flow sometimes involved a growing number of participants, branching out of ideas and insights and leading to a creation of new links between ideas.

The dialogic significance of this joint reflective process also concerns how the starting point of the statements of all three students is attentiveness and caring, and their acknowledgement that coping with the challenges of their collaborative learning is a significant issue that is worthy of their time and efforts. Furthermore, the way the three students express their self-criticism may testify to their appreciation of this learning community as a safe, protected space. However, as these exchanges are visible to comments from all community members, the meaning of the dialogic reflection is not limited to what is said by the three students. It also encompasses a view of this discursive event as part of the continuum of small discursive events in the overall context of the discourse taking place in the learning community, as they learn from and develop the ideas of one another.

Community Reflective Conversation as a Space for Coping with Interpersonal Conflicts

Although contrasting perspectives were often expressed during the CRC or in students' personal diaries, in most cases, it did not take the form of an argument, in the sense of an attempt to get to a single conclusion or to convince each other. Rather, in a tolerant and supportive atmosphere, disagreements were used to reveal new meanings and understandings. Many students considered such events as insightful and significant which is clearly evident in their personal diary reports. As an example of this, following an unsuccessful collaborative learning experience, one of the leading students used the next CRC to share her overall disappointment with the way the group handled that task. While criticizing her own part in that process as well, she raised doubts about the usefulness and the advantages of collaborative editing in general. This sharing evoked an emotional discussion about the social norms of the community in general and about the different roles students took during this specific event. Later on, the other leading student of that specific learning assignment wrote the following:

> Raida: I thought that Ronit and I have talked this issue thoroughly, and solved all disagreements between us. However, only today in the [community] reflection session, I have understood that I didn't really listen to her carefully enough before, not giving her the opportunity to tell me: Raida, you too didn't show any interest in others and provide them freedom in their work… After this experience I think we should add more norms, such as roles definition and schedule…

The CRC gave Raida and Ronit an opportunity to say things and to listen to each other in a way that apparently wasn't possible before. In choosing to open this issue up in front of all the community members, Ronit expressed her trust and confidence in them and in that forum. By doing this, she involved and engaged all the members in a discussion about taking responsibility over the learning processes. Raida too, did something similar. While based her comments on herself and Ronit, she moved to the community level, referring to "we" and suggesting that all group members take a step forward in taking more responsibility over learning. Raida was fully aware at that time that what she wrote was public.

Building on and Developing Each Other's Ideas

The following excerpt is a different example of how ideas may pass along the students:

> Rina: What is nice about Wiki—and this is an idea that I am developing a bit after hearing it from Ariel (who heard it from Shlomit)—is that the Wiki world is effectively a world with no borders…this sparked many questions in me: Does being dominant mean causing a kind of paralysis in someone else? Or is our goal as students in a collaborative community engaged in collaborative writing to find our place, to make our way, aspire to grow and develop as far as our abilities allow us, and at the same time be collaborative and attentive to others?

This excerpt illustrates a process of collaborative development of a theoretical idea. Moreover, Rina synthesizes an idea that she read in other student's personal diaries with a significant concept that appeared in the community reflective conversation in the context of a specific event. From this starting point, she continues to pose more general questions, and suggests answers that relate to the theoretical content learned in the course. This example shows the creative potential lies not only in the reflective writing in itself, but in the possibility of one student to connect with other's ideas.

A Growing Appreciation of the Community Dialogic-Reflective Discourse

The following excerpt illustrates a general approach to the learning experience and its significance:

> Michal: Some of the things I wrote or said attracted responses of community members, and I am very grateful to them for that. I think that the most significant learning that I did in this course was to recognize my ability to learn from others and grow from the dialogue taking place with community members following the ideas and thoughts that I expressed. This discovery was new, surprising, and addictive. I find that this dialogue allows me to reach places that I never even thought about, it can help break through dead-ends, it is enriching, and especially very interesting and fascinating.

This quotation is brought also to point to one of the most interesting developments we were witnessing: the gradual opening of many students to the possibilities of communicating and sharing with others within the community discourse. The importance of these quotations is that they show that the students themselves appreciate the dialogic aspects of the course, and recognize them as beneficial for themselves. On the one hand, reading other students' reflections can enrich the perspective of other students, trigger new thoughts and understandings and make them aware of issues and people in a new way. On the other hand, when other people respond to a student's reflections, that student is "being noticed" and "counted," which creates the feeling that she can make a difference to somebody else. Furthermore, being part of a dialogue that is considered significant and meaningful is to acknowledge learning as a social process that benefits its participants.

Discussion and Conclusions

In the following section, we discuss the dialogic and reflective dimensions of the CEDRD framework in relation to the design of the learning environment. We focus on the mutual exposure to the students' reflections, and the possibility for communication through varied and rich channels, as central features of the learning environment that enable dialogic-reflective discourse to emerge.

Students' mutual exposure to different reflections regarding the socio-emotional challenges of their collaborative interactions, provide them with a diverse picture of collaborative experience. This may enhance their awareness of the potential fragmentary nature of any single personal perspective. In this respect, the diverse picture is not necessarily a single, uniform understanding of an event, but rather the understanding that the event may have different interpretations that may be in tension with each other. Synthesizing individual perspectives is not a harmonious process that is free of challenges. Creating the big picture entails coping with dilemmas, conflicts, and criticisms that arise as part of the learning process. Therefore, we argue that it is specifically this coping process that represents the greatest potential of the dialogic-reflective discourse. It invites a constructive, critical, and collaborative negotiation of meaning of the emerging, diverse picture. This negotiation deepens students' understanding that each can contribute to and learn from the others, and promotes their appreciation of the learning community as a social framework that enhances and supports them.

The mutual exposure to reflections about theoretical knowledge studied in the course and their relevance to students' own lives, allows for a similar yet different process. In this case, it enables students to reconsider their original understandings or perceptions regarding this knowledge, and be enriched by and contribute to the insights of others. This also includes the view of the community as a space in which ideas and insights evolve and each student can and may develop or build upon insights of others. The exposure to students' reflections about their personal lives allows students to learn more about each other, in addition to their performances as students in the classroom. The emphasis here on mutual respect, care and responsibility that grows out of students' appreciation for each other. Closeness among learning community members does take place, eventually, as they grow to learn more about each other and engage together in the dialogic-reflective discourse.

Our discussion up until this point has concentrated on the significance of the mutual exposure to students' reflections in deepening and widening their acquaintance. Furthermore, we suggested that by this exposure they gain a richer perception of themselves as members in a collaborative learning community and of collaborative learning in general. This exposure has an additional importance which may be viewed as a door to a unique kind of dialogic communication among the learning community members: unique in terms of the themes of discussion as well as the manner in which the discussion is carried on. This communication may start with one student approaching the other with questions about a reflection. In other cases, this involves trying to get a clearer or deeper understanding of the other's point of view, and sometimes as a combination of both. Moreover, communication may start with the students questioning their own conduct, doubting if the way they acted was the only possible way to act. By doing that, instead of closing in while justifying one's own opinion or conduct, the discussants are opening up spaces within themselves and towards the other dialogically.

The open-to-all reflective and communicative nature of the CRC and the online personal diaries supports the acquaintance between students. It is a cyclic process: As students' acquaintance deepens, students

give deeper expressions to their emotions and insights, and their interest in and desire to respond to others' expressions also grow stronger. This acquaintance entails more than the knowledge of who a person is but the attempt to understand what they are going through (Noddings, 2012). This comes out of a sense of caring and attentiveness to the singularity of each student, and on the basis of the reflective expressions of each. The seeds of this potential are grounded in the teacher's dialogic approach and attitude, and this becomes a part of the students' own experience. An important element in this experience is the growing understanding and feeling that there is enough space for everyone, there is a place for everyone, and everyone counts. While that alone may carry a great significance, getting comments from others becomes so important because it shows that you have been noticed, and that your voice has been heard. Comments may also indicate that what you have said is meaningful to somebody, and at the same time it shows that somebody else cares about you.

Overall, the CEDRD framework allows us to identify the conditions under which collaborative learning develops into a meaningful process. Likewise, the framework demonstrated how a general appreciation of the contribution of collaborative learning to the students themselves emerges.

Limitations and Challenges

Emphasizing the potential significance and contribution of the CEDRD framework, we do not imply for the possibility of integrating this model in its fullness in each and every class or discipline. Rather, we want to suggest that it can be integrated in different levels or forms in accordance with the different conditions available. In pioneering research, Gofer (2013) showed how a more modest model may be successfully implemented with eighth graders. In her model, while students wrote their reflections in Wiki-based public diaries, conflicting incidents were not discussed in the whole class forum, but within sub-groups, under the guidance of the researcher and the teacher. In our case we are also fully aware of the fact that not all conflicts were discussed among students and not all personal dilemmas and difficulties were raised and shared while engaging in the community discourse. Moreover, there were always some differences between students, regarding their level and depth of sharing and their actual participation in the collaborative effort to face conflicts or to share and discuss theoretical insights. Bearing this in mind, preliminary results indicate that most students did engage in the community discourse, as well as the learning communities were stable enough to contain the ones that stayed less involved.

Summary

This paper presents the potential of dialogic-reflective perspective to contribute to our understanding of the socio-emotional aspects in collaborative learning communities. Furthermore, we suggested that the CEDRD framework may allow us to propose possible conditions under which students are engaged deeply in the collaborative learning processes and acquire a general positive perspective of collaborative learning as learning that enhances and empowers them. If we want students to deeply understand the meaning of collaborative learning and adopt it as *an approach to life*, they need to experience and examine it. We stress that this learning may be enhanced by students' engagement in a discourse that allows them to jointly confront and cope with the conflicts and dilemmas that arise from collaborative learning, discuss the significance of their collaborative learning, and discuss the contribution of the learning community's discourse itself for them. Furthermore, a productive and empowering discourse requires close relationships that are based on acquaintance, attentiveness and care among the learning community.

References

Alpert, B. (2008). Nodding's dialogue: Relationships and dialogue as the essence and path in education for care (Hebrew). In N. Aloni (Ed.), *Empowering dialogues in humanistic education* (pp. 236–258). Tel-Aviv: Hakibbutz Hameuchad.

Baker, M., Andriessen, J., & Järvelä, S. (2013). Introduction: visions of learning together. In M. Baker, J. Andriessen & S. Järvelä (Eds.), *Affective learning together: The socio-emotional turn in collaborative learning* (pp. 1-30). Routledge.

Baker, M. J., & Lund, K. (1997). Promoting reflective interactions in a computer-supported collaborative learning environment. *Journal of Computer Assisted Learning*, *13*(3), 175-193.

Ben-Zvi, D. (2007). Using wiki to promote collaborative learning in statistics education. *Technology Innovations in Statistics Education*, *1*(1). Retrieved from: http://escholarship.org/uc/item/6jv107c7.

Bengtsson, J. (1995). What is Reflection? On reflection in the teaching profession and teacher education. *Teachers and Teaching*, *1*(1), 23-32.

Bergman, S. H. (2007). Introduction (Hebrew). In M. M. Buber, *The dialogue on man and being* (pp. 11-48). Jerusalem, Israel: Bialik Institute.

Bielaczyc, K. (2006). Designing social infrastructure: The challenge of building computer-supported learning communities. *Journal of the Learning Sciences*, *15*(3), 301-329.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229-272). Cambridge, UK: The MIT Press.

Buber, M. (2007). *The dialogue on man and being* (Hebrew). Jerusalem: Bialik Institute.

Burbules, N. C. (1993). *Dialogue in teaching: Theory and practice*. New York, NY: Teachers College.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support* CSCL? (pp. 61-91). Heerlen: Open Universiteit Nederland.

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research in computer-supported collaborative learning: From design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning: Principles and products* (pp. 3-19). Netherlands: Springer.

Friedman, M. S. (1956). Martin Buber's philosophy of education. *Educational Theory*, *6*(2), 95-104.

Friedman, M. S. (2002). *Martin Buber: The Life of Dialogue*. New York: Routledge.

Gofer, R., & Ben-Zvi, D. (2014). Metacognitive aspects of writing personal-reflective diaries in a wiki-enhanced collaborative learning environment. In Y. Eshet-Alkalai, A. Caspi, N. Geri, Y. Kalman, V. Silber-Varod & Y. Yair (Eds.), *Proceedings of the 9th Chais Conference* (pp. 48-55). Raanana: The Open University.

Guzdial, M., Hmelo, C., Hubscher, R., Newstetter, W., Puntambekar, S., Shabo, A., Turns, J., & Kolodner, J. (1997). Integrating and guiding collaboration: Lessons learned in computer-supported collaboration learning research at Georgia Tech. *Proceedings of Computer-Support for Collaborative Learning (CSCL)* (pp. 91-100). Toronto, Canada: University of Toronto.

Hatton N., & Smith D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33-49.

Hod, Y., & Ben-Zvi, D. (2013). An adapted group psychotherapy framework for teaching and learning about CSCL. In N. Rummel, M. Kapur, M. Nathan & S. Puntambekar (Eds.), *Proceedings of the Tenth Computer Supported Collaborative Learning Conference (CSCL10)* (Vol. I, pp. 240-247). Madison, WI: University of Wisconsin.

Kerr, N., & Bruun, S. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology*, *44*(1), 78.

Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to learning in settings of collaboration. *Proceedings of the Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 308-313). Mahwah, NJ: Lawrence Erlbaum.

Kramer, K. P. (2013). *Learning through dialogue: The relevance of Martin Buber's classroom*. Plymouth, UK: Rowman & Littlefield.

Noddings, N. (1995). *Philosophy of education*. Boulder, CO: Westview Press.

Noddings, N. (2007). Caring as relation and virtue in teaching. In P. S. Ivanhoe & R. Walker (Eds.), *Working virtue: Virtue ethics and contemporary moral problems* (pp. 41-60). Oxford, UK: Oxford University Press.

Noddings, N. (2012). The caring relation in teaching. *Oxford Review of Education*, *38*(6), 771-781.

Salomon, G. (2000). *Technology and education in the age of information* (Hebrew). Haifa and Tel Aviv, Israel: University of Haifa and Zmora-Bitan publishers.

Salmon, G. (2002). Mirror, mirror, on my screen…; Exploring online reflections. *British Journal of Educational Technology*, *33*(4), 379-391.

Shner, M. (2012). *Born virtual: The free spirit of man in a borderless world* (Hebrew). Tel Aviv: Mofet Institute.

Wegerif, R. A. (2007). *Dialogic, education and technology: Expanding the space of learning*. New York, NY: Kluwer-Springer.

Acknowledgments

We thank the *Learning In a NetworKed Society Research Center* (LINKS I-CORE) for funding this study. We also wish to thank Mosheh Shner, Arieh Kizel and Tsafrir Goldberg for their important remarks. Special thanks to Yotam Hod and the members of the *COOL-CONNECTIONS* research group for their careful review of this paper.

A Case Study Examining the Microdynamics of Social Positioning within the Context of Collaborative Group Work

Lesley Dookie, University of Toronto, 252 Bloor St. W., Toronto, Canada, lesley.dookie@utoronto.ca

Abstract: During collaborative group work, students from non-dominant social groups can be positioned by classmates in ways that hinder their opportunities to learn and become successful mathematics students. Drawing from an episode of videotaped collaborative group work, this qualitative case study examines the microdynamics of positioning and, using a stimulated recall interview technique, explores how a girl who was working with a group of boys identifies, interprets, and explains these moment-to-moment acts of positioning. The findings point to the strength of this methodological triangulation by further elucidating verbal and non-verbal forms of positioning. Specifically, the results illustrate how the focal student was prevented access to shared learning artifacts and group discussion due to her group members' (likely unintended) 'exclusive talk' and 'physical blocks'. Whether and how the observed acts of positioning are associated with social categories (i.e., gender) are discussed and implications for the implementation of collaborative learning activities are raised.

Becoming successful in mathematics class is a challenging feat that is more attainable for some students than others. There is a growing body of work arguing that *who* students are influences their access to learning opportunities and the quality of their interpersonal interactions within the classroom (Gee, 2000). That is, women and students of color are negatively stereotyped and historically marginalized within the mathematics domain and research documents the ways in which mathematics learning is racialized and gendered (Esmonde & Langer-Osuna, 2012; Langer-Osuna, 2011; Martin, 2006). Within mathematics classrooms, collaborative learning activities wherein girls and/or students of color work with students from dominant social groups have been documented as being rife with issues of power and inequity (e.g., Esmonde & Dookie, 2012; Langer-Osuna, 2011; Leander, 2004). Given the trending reform movements in mathematics education that endorse meaning-making through collaborative learning opportunities (e.g., Gutiérrez, 2002), there is a push for equity focused research to understand not only the sociohistorical systems of power that underlie these collaborative learning contexts, but also how this power is constructed through moment-to-moment interactions between students. This understanding can help further elucidate the subtle yet pervasive ways in which students from non-dominant social groups are hindered from becoming successful in mathematics class.

Theoretical Perspectives

Understanding Learning and Identity

This work is grounded in sociocultural conceptualizations of learning and identity. Learning, according to sociocultural theory, is mediated by interpersonal interactions and the use of artifacts (i.e., objects that have become meaningful over time as a result of their repeated use in goal-directed human activity, Cole, 1996; Vygotsky, 1986). Sociocultural theorists link the process of learning with identity development (Wortham, 2006), positing that learning is not merely about acquiring knowledge and mastering skills, but includes shifts in identity (Lave & Wenger, 1991) and 'becoming' (Nasir, 2002).

From this theoretical standpoint, identity is a social construct that is fluid in nature (Nasir, 2002), emerges through interpersonal interactions (Holland, Lachiocotte Jr., Skinner & Cain, 2001), and evolves as one constucts a notion of self within the context of particular practices such as mathematics learning (Nasir & Hand, 2008). Nasir and Hand (2008) use the term *practice-linked identity* to describe, "the identities that people come to take on, construct, and embrace that are linked to participation in particular social and cultural practices" (Nasir & Hand, 2008, p. 147). There is an imaginative component to this process of identitiy development in that, over time and repeated experiences, students in mathematics classrooms, for example, may come to see themselves in ways that incorporate mathematics (Lave & Wenger, 1991; Nasir, 2002). For example, drawing from the principles of learning and identity developed by Nasir (2002), when a student comes to learn a new mathematical concept, this can offer them new ways to participate in collaborative activities which can allow them to further develop their identities relative to the mathematics community. By providing opportunities for meaningful participation, engagement, and self-expression, practices can support the development of positive practice-linked identities and thus faciliate learning (Nasir & Hand, 2008). At the same time, however, experiences that deny these opportunities and limit students to marginal forms of participation can lead to the development of negative practice-linked identities and hinder learning. With this in mind, the present study seeks to better understand the social processes that serve to marginalize students within the context of collaborative group work.

To account for the ways positive or negative practice-linked identities can develop over time through moment-to-moment interpersonal interactions, I draw on the notion of *positional identities*. Holland and her colleagues (2001) assert that moment-to-moment interpersonal interactions position individuals in the social world in ways that provide them with differential access to spaces, conversations, and overall participation in a practice. Moment-to-moment acts of positioning can be implicit or explicit and intentional or unintentional. To demonstrate the microdynamics of positioning and how it unfolds through interaction, Leander (2002) analyzed talk and traced the physical organization of students, including their eye gaze and bodily orientation. Through this microanalysis, Leander was able to demonstrate how one girl was 'silenced' by a group of boys during a classroom interaction. Within the context of collaborative mathematical group work, the ways in which students are positioned through their moment-to-moment interactions with classmates can have a profound impact on their opportunities to learn. For example, Esmonde and Dookie (2012) engaged in a microanalysis of student interactions within the context of collaborative group work to demonstrate the mechanics of marginalization and the ways in which one student was negatively positioned by her group members through both verbal and non-verbal means. Similarly, through the application of their *Differential Influence* framework, Engle, Langer-Osuna and McKinney de Royston (2008) demonstrated how students become influential in collaborative group work in part through their differential access to interactional spaces and the conversational floor. Dookie and Esmonde (2012) build on this work by demonstrating how the construction of power and influence is also shaped by students' access to shared learning artifacts within the context of collaborative group work, with greater access to artifacts facilitating more central forms of participation in the group activity.

Taken together, there is a growing body of work investigating the construction of power and the mechanics of positioning within collaborative learning contexts. This research is also beginning to reveal the ways in which positioning is tied to *social identity* (i.e., social categories such as race and gender). For example, empirical studies demonstrate how students from dominant racial and gender groups have greater access to learning opportunities within collaborative group work while the participation of students from non-dominant social groups is constrained (Langer-Osuna, 2011; Kurth, Anderson, & Palinscar, 2002).

Student Voice as a Critical Research Perspective

What is less understood is how moment-to-moment acts of positioning are experienced and interpreted by students from non-dominant racial and gender groups. Based on the tenets of critical race theory (Ladson-Billings & Tate, 1995), it is important to employ interview methods in research involving individuals from non-dominant social groups in order to provide them with voice. This perspective is essential in truly understanding the lived experiences of those who are marginalized and oppressed by society and is particularly important within the context of group work among culturally and racially heterogeneous students wherein marginalization may be routinely experienced by students from non-dominant social groups (e.g., Kurth et al., 2002; Langer-Osuna, 2011). Furthermore, in the endeavor to better understand how acts of positioning are associated with social identity, it is important to gain an understanding of how these acts are interpreted by those who are on the receiving end. For example, *microaggressions* (i.e., "brief and commonplace daily verbal, behavioral, and environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative racial slights and insults to the target person or group" Sue, Capodilupo, Torino, Bucceri, Holder, Nadal, & Esquilin, 2007, p. 273) can be conceptualized as forms of marginalization that are tied to particular social categories. Microaggressions manifest in everyday interracial and/or intersexual interactions and are subtly delivered verbally and non-verbally. Particular facial expressions, gestures, tones, and utterances can be microaggressions depending on the context (Solórzano, Ceja, & Yosso, 2000). For example, within a workplace, if a male addresses a female subordinate as 'sweetie', she may experience this interaction as a microaggression. Researchers who investigate microaggressions assert that these marginalizing experiences are subjective in nature, can only be identified by those who experience them, and are therefore best investigated through interview methods (Solórzano et al., 2000; Sue et al., 2007).

Previous research investigating the mechanics of positioning involves video recorded observations of student collaboration as well as microanalyses of this data. This work is limited to only what can be seen in video data. As third party observers, there is danger in making assumptions about students' interpretations and experiences (e.g., Leander, 2004). The present study aims to extend previous work by adding a layer of analysis that considers a focal student's reaction to verbal and non-verbal forms of positioning. Specifically, focusing in on a segment of video recorded collaborative group work and using stimulated recall interview techniques, this case study seeks to investigate: (1) the microdynamics of positioning within the context of collaborative group work and (2) whether and how a focal student from a non-dominant social group identifies, interprets, and explains these acts of positioning. This methodological triangulation can offer further insight into the microdynamics of positioning and provide a more nuanced understanding of how students from non-dominant social groups experience collaborative group work.

Methodology

Participants and Research Context

The data for this qualitative case study were collected as part of a larger study investigating the role of social identity in mathematical group work. This case study is focused on the experiences of Heather (all names are pseudonyms), a multiracial girl, as she engaged in a collaborative learning activity with three White boys. Heather was 16 years old and enrolled in an 11th grade Advanced Functions mathematics class at a private school in a Canadian city. Heather identified herself as a White/Asian girl. The analysis is focused on a session of group work wherein Heather worked with Allen, Walter, and Paul on a set of word problems involving the use of advanced functions. The students were given a shared worksheet listing the three word problems as well as access to a whiteboard to record their calculations and solutions. The group membership was assigned by the classroom teacher and the students were encouraged to collaborate. For the purposes of this case study, the analysis was focused on a segment of video wherein the students were solving a word problem involving statistical probability and the use of a Venn diagram.

Methods and Data Sources

The group work activity was videotaped and brief fieldnotes were taken to complement the video footage. Immediately following the observation, the video was processed and a minute-by-minute content log of the footage was created. The fieldnotes and video were used to identify five moments of interest in the group work activity, including times when Heather appeared to be marginalized (i.e., prevented access to the conversational floor and/or shared learning artifacts) or during intergroup interactions that were deemed interesting by the researcher.

A day following the group work observation, an audio-recorded stimulated recall interview (SRI) was conducted with Heather. Using the video footage of the group work as a prompt, she was invited to watch the video, stop it at any time, and share her reactions, interpretations, feelings and so forth. If Heather did not independently stop the video to comment, the researcher did so after the segment ended and invited Heather to respond to the footage. Heather watched the five segments of video that were pre-selected by the researcher. These video segments ranged in length from about 30 seconds to 2 minutes and 30 seconds.

Approximately one week following the SRI, Heather was interviewed a final time to determine her more general feelings and experiences with group work as well as her knowledge of and experiences with mathematical achievement stereotypes. Heather was encouraged to draw on personal experiences, including the observed episode of group work.

Analysis

A *context analysis* (Erickson, 2006) focusing on talk, gesture, body positioning, and artifacts (i.e., the shared worksheet and whiteboard) was conducted on the five video segments of interest using Studiocode video analysis software. Drawing from Erickson's (2006) *whole-to-part* procedure for analyzing video, each segment of interest was played in its entirety, multiple times with and without sound. To further section the video, major transitions in the bodily configuration of students was marked on a video timeline. Each of these smaller sections of video was then analyzed at a micro-level with a specific focus on both verbal and physical forms of positioning in relation to Heather. These instances were marked on the video timeline and, drawing from the analysis of Esmonde and Dookie (2012), Heather's subsequent verbal and non-verbal actions were coded. After an initial pass of marking these instances of interest on the video timeline, a system of codes was developed. The development of this coding system was an iterative process involving repeated viewing of the video segments, multiple passes of coding, and the refinement of codes.

A full description of the codebook is beyond the scope of this paper, however, in general, the codes were used to identify verbal and non-verbal acts of positioning in relation to Heather as well as her observed responses. Particular attention was paid to the participants' bodily orientations and movements in space relative to Heather and whether/how they impacted her access to the interactional space and the learning artifacts. A transcription of the students' talk was used to analyze whether and how Heather was given opportunities to participate in the group discussion (note that in the selected segment of video presented in this case study, Heather did not make any utterances and hence her talk was not part of this analysis). Examples of positioning codes used in this analysis include 'physical block' (any time a group member physically blocked Heather's access to the worksheet or whiteboard) and 'exclusive talk' (any time members of the interaction physically oriented themselves to talk to another member of the group or verbally addressed a group member by name to participate in the activity). An example of a code that was used to characterize Heather's 'responses' to acts of positioning includes 'strain' (any time Heather had to lean or contort herself, while keeping the trunk of her body stationary, to better see and access the shared artifacts). Codes such as 'access to worksheet' and 'writing' were used to identify Heather's access to the shared learning artifacts.

The Studiocode software facilitated an analysis of the frequency of codes used as well as the timing in which they were applied. This information was used to write an analytic memo for each of the five video segments of interest. In the results section below, I will share the detailed findings pertaining to one of the video segments of interest as well as Heather's response to this video segment. The video segment was 2 minutes and 30 seconds in length and was selected because it was representative of the group work episode and depicted several examples of both verbal and non-verbal forms of positioning.

Together, the interviews and video data informed one another. In addition to providing the participant perspective and voice, the interviews provided further context to the episode of group work. Drawing from Anderson (2009), a limitation of traditional analyses of positioning is that they employ an *imminentist ontology* (i.e., "the premise that positioning is contextually tied to the moment of interaction in which it occurs and not across interactions or scales of activity" (p. 292) and she calls for the need to consider meso-level influences (i.e., "neither from a solely micro- nor a solely macro-social perspective"; p. 293). As such, in addition to considering the moment-to-moment acts of positioning and broad sociohistorical forces such as gender, this study used the interview data to facilitate the investigation of some the meso-level influences (e.g., Heather's relationship with the group members) that came to bear on the episode of group work.

The interviews were transcribed and exposed to descriptive followed by thematic passes of coding. While talk involving social identity, marginalization, and stereotypes was a primary focus, themes (e.g., friendship, dispositional characteristics of group members) also emerged from constant comparison between data and the gradual elaboration of open codes (Glaser & Strauss, 1967). An analytic memo was written for both interviews and the themes were elaborated.

Results

To begin, a vignette is presented to illustrate the ways in which Heather was physically and verbally marginalized in the group work activity. The microanalysis of the selected video segment is then juxtaposed with the narrative interpretation of the experience provided by Heather after she viewed the same segment of video and engaged in the SRI. Following this micro-level discussion about the acts of positioning, Heather's interpretations and feelings about the experience are reviewed.

Part I. Microanalysis of a Segment of Group Work

Vignette: Observed Positioning

Heather, Allen, Walter, and Paul assembled as a group and began the first of three word problems. Allen held the worksheet and he and the other members of the group physically oriented themselves towards the shared worksheet as well as the whiteboard (see Figure 1, image on the left). The initial configuration of the group was semi-circular with Heather slightly on the outskirts. This physical position left her with only moderate access to the worksheet meaning she would have had to significantly lean and contort herself to read the word problem. She had relatively clear access to the whiteboard space wherein the problem was being solved.

Allen began the collaborative group work by reading the word problem aloud. He and Walter discussed the given information and how they would construct the Venn diagram while Paul and Heather watched on. Allen instructed Paul to begin drawing the Venn diagram and from this moment onward, each time Paul went to write something on the whiteboard, he created a physical barrier that made it more challenging for Heather to access the whiteboard (i.e., the central space for the group work activity; see Figure 1, image on the right).



Figure 1.The left image illustrates the original configuration of the group, depicting, from left to right: Allen, Walter, Heather, and Paul. The right image illustrates how Paul blocks Heather each time he writes on the board.

Allen and Walter continued to reread the problem aloud as Paul filled in the given information on the Venn diagram. Once the given information was plotted and the discussion about how to solve the problem became the focus, Walter gradually positioned himself more square in line with the whiteboard and thereby further blocked Heather from the center of the group work activity. Together, the body positioning and physical movements of Paul and Walter served to physically box Heather out of the activity (see Figure 2, image on the left).

Towards the latter half of the video segment, once the physical blocking became most prominent, Heather appeared to strain herself (i.e., lean in, angle her head) to better see the worksheet and what was being

written on the board. After a few moments, Heather walked to a new position in the group wherein she was closer to the whiteboard yet further from the worksheet and the intermittent discussion between Allen and Walter (see Figure 2, image on the right).

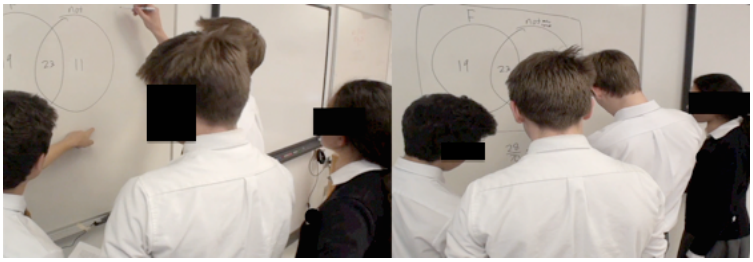


Figure 2. The image on the left depicts the physical blocking by both Walter and Paul. The image on the right depicts Heather's new spatial position in the group.

The microanalysis of body positioning during this video segment revealed that Heather was on the physical periphery of the activity with limited access to the interactional space. As the physical barriers created by her group members became more prominent and stable, she strained herself to better see the information on the worksheet as well as what was being written on the whiteboard and eventually she moved to a new spatial location in the group. Although this move positioned Heather closer in proximity to the whiteboard, it further distanced her from the center of the group discussion (i.e. between Allen and Walter) and the shared worksheet.

During this video segment, Allen was the only person to hold the worksheet and Paul was the only member of the group to write on the board. In the analysis of talk, five clusters of conversations were identified, primarily involving either Allen, Walter, and Paul or only Allen and Walter. Furthermore, three out of five of these conversations were characterized as 'exclusive talk', meaning that members of the conversation either explicitly addressed one another in the conversation or physically positioned themselves towards another person thereby excluding other(s) (i.e., mainly Heather) from participation. Heather did not make a single utterance during this video segment (and said very little as they solved the subsequent two word problems).

From the theoretical and analytical lens of the researcher, this microanalysis reveals the ways in which Heather was negatively positioned by her group members. The question remains as to whether these physical blocks and exclusive conversations between the male members of the group were a salient part of the experience for Heather. Let us now turn to the interview data to investigate whether and how Heather identified, interpreted, and explained these acts of positioning.

Part II. Using Interview Data to Explore Heather's Experience

Identifying Acts of Positioning

The video segment described in the previous section was played for Heather during the SRI and she was then asked to respond to the footage and describe what she was thinking and feeling during that point in the activity. She began by explaining that she did not enjoy the word problem because the people in her group were "taking charge". She went on to describe how she was physically blocked and had difficulty seeing the worksheet and the whiteboard:

> They didn't- I couldn't really see the question, so I didn't really *know* what was happening and then I was trying to look at the board and then [Paul] decides to like stand right there in front of me. So I'm like, "oh, ok!"

A few moments later, she went on to say:

> Well also like, [Paul] is like blocking my view so I can't really see what's going on. And they're just like...and those- [Allen] and [Walter] or- those people, they like-they like taking charge and that's what they're doing! …And also like, I didn't know what the question was cause I couldn't see it cause they're holding it [the worksheet] like *that*, so I didn't- couldn't really contribute anything.

Heather's narrative clearly illustrates the impact of physical marginalization. Here and throughout the SRI, she repeated the fact that she was being physically blocked by Paul. In contrast to the microanalysis of the video that revealed the physical blocks of both Walter and Paul, in her narrative, Heather solely focused on the blocking actions of Paul. To further build on the findings of the video microanalysis, Heather attributed her inability to access the worksheet as a result of the way in which Allen was holding the paper and angling it towards himself:

Heather: Just like- I don't really know what's going on cause I haven't heard the question.
Researcher: Yeah. Cause you didn't hear the question being read?
Heather: Yeah and like, he's [pointing to Allen] hoarding the question.

As a result of this limited access to the shared learning resources, Heather indicated that she "couldn't really contribute anything". Heather went on to explain her movement to the new spatial location in the group as "an attempt to see the board" as a result of being blocked by Paul. Consistent with the microanalysis of the video segment, Heather indicated that although the move allowed her to see the board better, "that was about it" and it did not increase her participation in the group activity.

In addition to discussing the physical aspect of the marginalization, Heather acknowledged the way Allen and Walter dominated the discussion and essentially took charge of the group work. She also noted that Paul did the writing for the group and served as the "scribe" for Allen and Walter. Taken together, Heather's narrative points to the subtle ways in which marginalization occurred through verbal and nonverbal means and provides further nuance into the experience of this marginalization.

Interpreting and Explaining Acts of Positioning

Drawing from the full episode of video recorded group work, the SRI, and the final interview, the focus of the analysis is now broadened to consider some of Heather's general interpretations and explanations. At the end of the SRI, Heather was asked to describe her general feelings about the group work activity:

> Well, I did not particularly enjoy it, but I think for people like [Allen] and [Walter]- like people who really dominate- that was helpful for them cause they could like bounce ideas off each other. But personally, I did not really benefit from that cause I couldn't see the questions and then I didn't really know what was going on and then…yeah.

Based on Heather's appraisal, the physical marginalization seemed to be a salient aspect of the overall group work experience. A microanalysis of video segments obtained across the entire group work footage revealed that the physical blocks, 'hoarding' of the worksheet, and exclusive conversations were not isolated events. Heather also explained that she was often grouped with these same students within the mathematics class and they characteristically tended to take charge of the group work. When asked how this group work experience could have been improved, Heather suggested having the group "work in a circle instead of a line" so that all members could see the worksheet and the whiteboard. Heather also advocated having students select their own group members: "I know like you're not supposed to work with your friends, but I actually would prefer working with my friends cause then they would actually let me *see* the question and then we could actually find out what's going on." Taken together, the importance of the physical dimension of the collaborative activity was emphasized throughout Heather's narrative description of the experience.

Although Heather identified various forms of marginalization and provided insight into how it made her feel, she did not offer as many clear explanations to account for these actions. During the SRI, she did, however, use the dispositional characteristics of her group members to account for their actions. For example, with respect to Allen 'hoarding' the worksheet she stated that, "he loves his paper" and described him as being "controlling with math". She also indicated that Allen and Walter were, "people who really dominate" and "enjoy taking charge" and so they tended to do so during the activity. During the final interview, when asked about whether race or gender had an impact on the group work dynamic, she stated:

> I don't think race comes into play. Especially not at [my school] cause [my school] is a very nice and diverse school. But maybe gender a little bit. Like I don't even think it's necessarily gender, it's more like who you're friends with. Cause the guys, they were all friends.

This quote reveals Heather's perception of the school culture as a "nice and diverse" place devoid of issues of race. It also suggests the underlying importance of friendship for Heather within group interactions. Although she only alludes to the role of gender in the group work interactions, her reference to the boys in the group all being friends suggests that these friendships are gendered in nature.

Discussion and Implications

This case study provides a snapshot of how one student became a marginal member of a group through the ways she was positioned by her group members. Whether or not Heather was proficient with the mathematical content or had ideas to contribute to the group, the exclusive exchange of talk that took place between her group members as well as the ways in which they oriented their bodies, prevented her from accessing the interactional space and conversational floor. Together, this limited her to peripheral participation in the activity. Reinforcing the work of Dookie and Esmonde (2012), this study also demonstrates the significance of shared learning artifacts in collaborative learning contexts. Not being able to see the whiteboard or access the shared worksheet

were central concerns highlighted across the analyses and contributed to the ways in which Heather was restricted to participating in the margins of the activity.

By considering both observed and experienced positioning, this case study makes a unique research contribution. It offers a nuanced understanding of how a student from a non-dominant social group experienced collaborative group work and provides further insight into the microdynamics of positioning. Although they foregrounded different dimensions of marginalization, the analysis of both the video and the SRI yielded insight into how Heather was marginalized through verbal and non-verbal means. For example, the findings of the microanalysis highlighted Heather's limited access to the shared worksheet while her narrative provided further detail and described that it was the way in which Allen was holding the worksheet that made it difficult for her to access it. The video footage alone (and the constraints associated with using a single video camera) could not provide this kind of detail and nuance. Similarly, while the findings of the video microanalysis emphasized the way Heather was physically blocked by multiple members of the group, Heather's narrative focused specifically on the physical blocks of one group member. These findings are similar to those of Leander (2004), who was surprised to find that although his analysis of a video recorded interaction was centered around student-student positioning, a posthoc interview with a focal student, Latanya, instead revealed her focus on her relationship to the classroom teachers (who were not captured in the video footage). Leander (2004, p. 207) states, "While Latanya and I watched the same videotape of the interaction, she appears to have attended primarily to [the classroom teachers] who were, for the most part, off camera." These findings point to the potential for misinterpretation when we rely too heavily on the sole perspective of researchers in these microanalyses of social interaction and highlight the importance of considering the perspective of the participants involved.

This study intentionally focused on the account of one focal student. It is important to acknowledge that reflective perspectives of an event may often differ from the complex ways in which actions are collaboratively organized in the moment. Positioning is an interactional dance that is co-constructed by the participants involved. This study focuses on the ways in which a student from a non-dominant gender group was restricted access to participation. Heather's moves in this interactional dance were limited by the moves of her group members. By straining her body and moving to a new physical location in the group, however, Heather made physical bids to be included. One could imagine a number of other ways in which Heather could have pushed back. Future studies should further examine the ways in which a person can be restricted from demanding access to central participation and whether and how this is tied to social identity.

It is also important to note that an interview provides a subjective perspective of an event and, in addition, is always a conversation between two people, rather than insight into all the participant's private thoughts. This is particularly important in reference to the interpretations and explanations provided by Heather in her response to the video footage. From the theoretical and analytical lens of the researcher, Heather was working with three boys and their physical blocks and exclusive talk could be conceptualized as microaggressions. However, Heather did not vocally explain these acts of positioning as such and essentially dismissed the notion of race and gender playing a role. Instead, her talk about gender was implicit in nature and came about through her indirect reference to the gendered nature of friendships within the classroom. Without undermining Heather's interpretation of the experience, this does not necessarily mean that gender and/or race were irrelevant in this collaborative learning activity. In her work examining *colormuteness*, Pollock (2004) demonstrates that there is a general propensity for students to resist talking about race in school in certain situations, despite its omnipresence. There are numerous reasons to account for why Heather may not have made explicit mention of taboo topics. For example, like many, Heather may have been motivated to appear 'egalitarian' which is particularly likely given the way she described the culture of the school as being "nice and diverse". One of the goals of this case study was to foreground the experience and voice of a student from a non-dominant social group and the intent here is not to discount her narrative but to instead draw attention to the subtle ways in which students may resort to talk about taboo issues as a result of school culture and so forth.

In her interview, Heather talked about gender implicitly by making an indirect reference to the gendered nature of friendships. In the endeavor to create more equitable learning contexts that provide students with opportunities to thrive and develop positive mathematics identities, it is important to understand the ways in which students experience and talk about their group work experiences. This includes the ways in which they talk about race and gender. Furthermore, it is important to investigate and understand the microdynamics of positioning so that educators can work towards better organizing collaborative learning activities so that all students have equitable access to learning opportunities. The classroom teacher in this case study had equity-focused intentions for the collaborative group work activity and it is likely that the acts of positioning performed by Heather's group members were unintentional and unconsciously delivered. However, students from non-dominant social groups can become marginalized in very subtle and unintended ways during collaborative learning activities. Over time and repeated experience, this can have a profound impact on their evolving mathematics identities. Bringing discussions about the microdynamics of positioning and marginalization into mainstream teacher discourse and professional development may be one way to work towards ensuring that all students have the opportunity to become successful mathematics students.

References

Anderson, K. T. (2009). Applying positioning theory to the analysis of classroom interactions: Mediating micro-identities, macro-kinds, and ideologies of knowing. *Linguistics and Education, 20*, 291-310. doi:10.1016/j.linged.2009.08.001

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press of Harvard University Press.

Dookie, L. & Esmonde, I. (2012). Understanding influence in collaborative group work: The importance of artifacts. In *Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)*.

Engle, R. A., Langer-Osuna, J.M. & McKinney de Royston, M. (2008). Toward a model of differential influence in discussions: Negotiating quality, authority, and access within a heated student debate. In B.C. Love, K. McRae & V.M. Slousky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2010-2015). Austin, TX: Cognitive Science Society.

Erickson, F. (2006). Definition and analysis of data from videotape: Some research procedures and their rationales. In J. L. Green, G. Camilli, & P. B. Elmore (with A. Skukauskaité & E. Grace) (Eds.), *Handbook of complementary methods in education research* (pp. 177–191). Hillsdale, NJ: Erlbaum.

Esmonde, I., & Dookie, L. (2012). The Mechanics of Marginalization and Resistance: Positioning in a Mathematics Classroom. Paper presented at the Research Pre-session of the Annual Meeting of the National Council of Teachers of Mathematics (NCTM 2012), Philadelphia, PA, USA.

Esmonde, I. & Langer-Osuna, J.M. (2012). Power in numbers: Student participation in mathematical discussions in heterogeneous spaces. *Journal for Research in Mathematics Education, 44*(1), 288-315.

Gee, J.P. (2000). Identity as an analytic lens for research in education. *Review of Research in Education, 25*, 99-125.

Glaser, B.G. & Strauss, A.L., (1967). *The discovery of grounded theory*. Chicago, IL: Aldane.

Gutiérrez, R. (2002). Beyond essentialism: The complexity of language in teaching mathematics to Latino/a students. *American Educational Research Journal, 39*(4), 1047–1088.

Holland, D., Lachiotte, W., Jr., Skinner, D., & Cain, C. (2001). *Identity and agency in cultural worlds.* Cambridge, MA: Harvard University Press.

Kurth, L. A., Anderson, C. W., & Palinscar, A. S. (2002). The case of Carla: Dilemmas of helping *all* students to understand science. *Science Education, 86,* 287–313. doi:10.1002/sce.10009

Ladson-Billings, G., & Tate, W. F. (1995). Toward a critical race theory of education. *Teachers College Record, 97*(1), 47–68.

Langer-Osuna, J.M. (2011). How Brianna Became Bossy and Kofi Came Out Smart: Understanding the Trajectories of Identity and Engagement for Two Group Leaders in a Projects-Based Mathematics Classroom. *Canadian Journal of Science, Mathematics, and Technology Education, 11*(3), 207-225.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Leander, K. M. (2002). Silencing in classroom interaction: Producing and relating social spaces. *Discourse Processes, 34*, 193–235. doi:10.1207/S15326950DP3402_4

Leander, K. (2004). They took out the wrong context: Uses of time-space in the practice of positioning. *ETHOS*, *32*, 188–213.

Martin, D. B. (2006). Mathematics learning and participation as racialized forms of experience: African American parents speak on the struggle for mathematics literacy. *Mathematical Thinking and Learning, 8*(3), 197–229.

Nasir, N.S. (2002). Identity, goals, and learning: Mathematics in cultural practice. *Mathematical Thinking and Learning, 4*(2-3), 213-247.

Nasir, N., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences, 17*(2), 143-179. doi: 10.1080/10508400801986082

Pollock, M. (2004). *Colormute: Race talk dilemmas in an American school*. Princeton, New Jersey: Princeton University Press.

Solórzano, D., Ceja, M., & Yosso, T. (2000). Critical race theory, racial microaggressions, and campus racial climate: The experiences of African American college students. *Journal of Negro Education, 69*(1/2), 60-73.

Sue, D.W., Capodilupo, C.M., Torino, G.C., Bucceri, J.M., Holder, A.M.B., Nadal, K.L., & Esquilin, M. (2007). Racial microaggressions in everyday life. *American Psychologist, 62*(4), 271-286. doi:10.1037/0003-066X.62.4.271

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans.). Cambridge, MA: MIT Press.

Wortham, S.E.F. (2006). *Learning identity: The joint emergence of social identification and academic learning*. New York, NY: Cambridge University Press.

Showing What They Know: Multimedia Artifacts to Assess Learner Understanding

Cindy E. Hmelo-Silver, Indiana University, 1900 East 10th St, Bloomington, IN, 47406 USA
chmelosi@indiana.edu
Carolyn A. Maher, Marjory F. Palius, Robert Sigley, Alice Alston, Rutgers University, 10 Seminary Pl., New Brunswick, NJ 08901, USA
carolyn.maher@gse.rutgers.edu, marjory.palius@gse.rutgers.edu, robert.sigley@gse.rutgers.edu, alston@rci.rutgers.edu

Abstract: Engaging learners in constructing multimedia artifacts provides rich opportunities for them to make their thinking visible. In this research, we demonstrate the use of the VMCAnalytic, a multimedia artifact that builds on an extensive video collection of children's mathematical reasoning. Using reliable rubrics, we coded all VMCAnalytics created in a range of classes. These rubrics focused on the quality of the students' arguments and depth of their reasoning. Analysis showed that the rubric was useful in differentiating among the different groups of students. Moreover, different metrics had different degrees of correlation, suggesting that we were identifying several different dimensions of quality.

Engaging students in technology-rich projects provides opportunities for both learning and making their thinking visible (Collins & Halverson, 2009). Creating multimedia artifacts offers opportunities for learners to engage with substantive content through their designs (Kafa & Ching, 2001). Building on earlier research with a video repository, we provide opportunities for students to engage in generative activity through construction of multimedia artifacts by making use of a new tool, the VMCAnalytic (Agnew, Mills, & Maher, 2010). In prior work, we examined a range of course contexts and tasks in which learners used the VMCAnalytic (Hmelo-Silver et al., 2013). In current research, we extend the range of contexts in which learners' use of the tool to construct multimedia artifacts enables assessment of the complex knowledge required for understanding, teaching, and researching the development of mathematical reasoning in students across several content domains. In particular, we examine how graduate students create arguments using videos of student reasoning by bringing together ideas from mathematics education and the learning sciences with the perceptual grounding of classroom practice to warrant claims about learning. Our research questions for this paper are as follows:

(1) To what extent can we use a cyber-enabled multimedia construction tool to assess how well learners justify their arguments about children's reasoning?
(2) To what extent do students identify relevant concepts in making their claims?
(3) How, if at all, does variation in course context, with differing instructional guidelines for completing a task, relate to qualitative differences in the multimedia artifacts that students produce?

We conjecture that studying these questions will guide the further development of formative and summative assessments. Conducting such investigations is of particular interest to the field of the learning sciences, where practitioners often fill the dual roles of designing activities for student learning and assessing the effects of those activities in order to learn what works to enhance student learning (Schwartz & Hartman, 2007).

The complex knowledge that we expect learners to construct through working with video, text, and practical experiences, entails multiple classes of learning outcomes. Included in their learning space is seeing, saying, and engaging ideas as well as targeting discernment, explanations, and contextualization (Schwartz & Hartman, 2007). Construction of multimedia artifacts with the VMCAnalytic tool supports student assessment by allowing us to identify the strength of arguments posed and the quality of their reasoning about mathematics education and the learning sciences.

Our research investigates the affordances and constraints of using multimedia artifact construction as a means of assessing the knowledge that learners have constructed about the development of mathematical reasoning and what they view as implications for teaching, learning, and/or further research. First we discuss the theories of learning that framed the earlier research yielding the video collection on children's mathematical reasoning and the recent research on teachers studying those videos to attend to students' reasoning. This is fundamental and relevant to the theory of learning through artifact design as a context for the current research. We then describe the resources on the video repository to illustrate what learners have available for constructing their multimedia artifacts, and share some results demonstrating the promise of the videos for learning about students' mathematical reasoning. This serves as a basis for exploring a technology-enhanced assessment for learners to show us what they know.

Theoretical Perspectives

Teachers seeking to promote students' competency to represent, communicate, and justify their ideas in the context of doing mathematics are faced with the challenge of developing their own adaptive expertise as educators (Bransford, Derry, Berliner, Hamerness, & Darling-Hammond, 2005*)*. High quality teaching demands the ability to spontaneously and flexibly identify, critically evaluate, and respond in appropriate ways to instances of children's learning. In mathematics it is particularly important to attend to emergent forms of reasoning as children express justifications using their own language (Hiebert et al., 1997; Yackel & Hanna, 2003). To build such capacity, teachers must know how to solve math problems, but must also come to recognize and understand the reasoning that justifies valid solutions to those problems (Maher, Landis & Palius, 2010). Consistent with the view of active knowledge construction, teachers need opportunities to engage as learners in building knowledge for teaching mathematics. There are several models for mathematics teacher education and professional development that use video as a tool to make instructional practices available for study, interpretation and discussion (e.g., Borko, Jacobs, Eiteljborg, & Pittman, 2008; Zhang, Lundeberg, & Eberhardt, 2011). The VMC video collection offers particularly valuable resources for teachers to build understanding of how students learn mathematics and conditions that promote development of mathematical reasoning. Because teachers typically have learned math in ways that relied heavily on procedural knowledge, facilitation of their learning entails engaging them in problem-solving tasks and justifying their solutions to build a deeper conceptual understanding of the mathematics and learning to attend to ways that students engage in those activities by studying video episodes.

Our goal is to advance teacher learning to the next level by engaging them in the construction of multimedia artifacts for sharing what they have come to understand about the development of mathematical reasoning or about learning more broadly. We define artifacts as "digital representations created by students that communicate their understanding, application, analysis, or evaluation of relevant ideas." (Rodriguez, Frey, Dawson, Liu, & Ritzhaupt, 2012, p. 358). While teachers' prior learning activities involved a video playback tool, their work with VMCAnalytic utilizes a video-editing tool. As with other video editing tools, such as WebDiver (Zahn et al., 2010), the VMCAnalytic enables selection of segments of video that can be annotated and remixed to form multimedia narratives of reflection and analysis for a variety of purposes (Hmelo-Silver et al., 2013). VMCAnalytics are shared, as they become objects of discussion, whether created individually or collaboratively. The process of constructing the analytics allows meaning to be negotiated and potentially refined for greater clarity and coherence.

Using a Video Repository

Support from four National Science Foundation grants for longitudinal and cross-sectional research studies produced over 4500 hours of video and related data showing students doing mathematics from elementary grades throughout high school and beyond. The Video Mosaic Collaborative (VMC, see: www.videomosaic.org) was built as a repository that houses a unique video collection, amassed from a quarter of a century of research (Maher, 2008; Agnew et al., 2010). An important early finding from this research is that young children, in justifying their solutions to problem tasks, provide arguments that take the form of mathematical proof (e.g., analogy, cases, contradiction, induction, upper/lower bound, Maher & Davis, 1995; Maher & Martino, 1996,). From our collection, one can study videos that show the evolution of the development of students' arguments over several years and follow how students' ideas are originally represented, shared, then expanded, and then generalized. The videos give examples of children's early strategies and heuristics and the durability of their early ideas, later expressed in more elegant form. One can follow the elaboration of student reasoning over several years and observe the richness and depth of understanding that has evolved over time as students make connections between and among ideas and express their solutions using formal notation and language to give meaning to the symbols (Maher, Powell, & Uptegrove, 2010). Because the research was conducted in working class, urban and suburban environments and in classrooms as well as informal, after-school settings, the collection is rich, not only from its longitudinal nature (some students are followed throughout schooling and beyond) but also because of the variety of contexts and content strands in which the research was conducted. As part of ongoing work, we continue to populate the VMC with series of short clips and full-length videos of problem-solving sessions from the collection by cataloging them with extensive metadata to enable a variety of search paths for discovering resources.

As videos get ingested to the VMC, they become publicly available resources for use in teaching, teacher education, and research (Hmelo-Silver et al., 2013). They also become available for use in the workspace of the VMCAnalytic tool as Figure 1 shows (Agnew et al., 2010). Following initial piloting and design revisions, we engaged learners in using the VMCAnalytic tool to construct multimedia artifacts for a range of purposes in a variety of graduate course and research contexts. Experiences with the VMCAnalytic tool have been preceded by participation in an instructional intervention during which videos from the collection were studied.

We have engaged in a program of design-based research to use the VMC resources for teacher education studies (Bielaczyc, in press). In particular, for pre and in-service teacher interventions, we have examined the effect of studying videos of student reasoning on teachers' growth in recognizing forms of reasoning used by children in the videos. Interventions were conducted in which study participants were asked to describe the details of children's arguments offered in justifying solutions to a specific problem-solving tasks (Palius & Maher, 2013; (Maher, Palius, Maher, Hmelo-Silver, & Sigley, 2014). In each strand, the task elicited multiple forms of reasoning that participants could recognize.

In the counting strand, a study based on our initial interventions investigated whether teacher study of VMC videos improved their ability to recognize a variety of forms of reasoning expressed by the children in the assessment video (Maher, Palius, Maher, Hmelo-Silver, & Sigley, 2014). Video assessment data were analyzed to measure growth from pre to post assessment for recognition of forms of reasoning. Results indicated that, on average, 60% of the in-service teachers and 36% of the experimental pre-service teachers improved on the post-assessment in recognizing the various forms of children's arguments, compared to the average of 5% improvement for the comparison group of pre-service teachers.

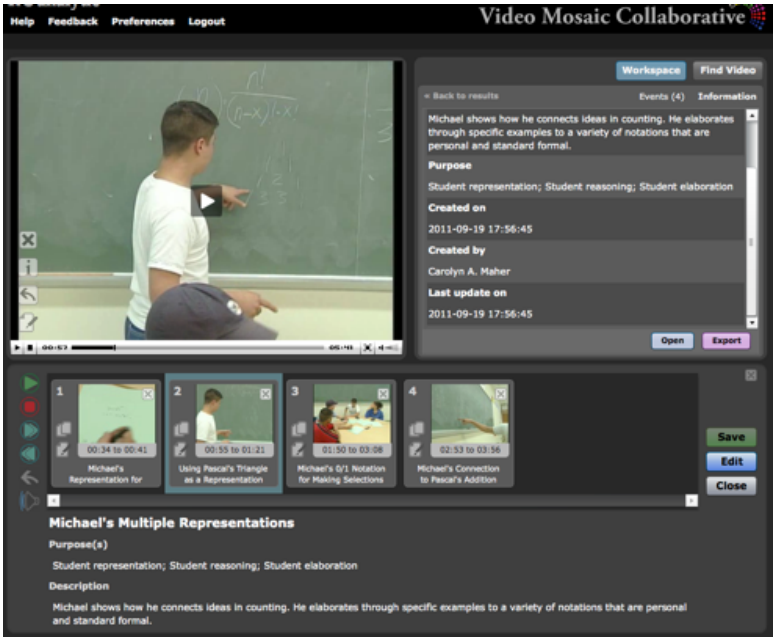


Figure 1. Screenshot of the VMCAnalytic tool

A study conducted in the fractions strand investigated teacher learning in an experimental online course using discourse analysis as well as analysis of their video-based assessment data (Palius, 2013). Teacher-learners in the experimental intervention were more likely to demonstrate growth in their ability to recognize different forms of students' mathematical reasoning compared with comparison groups (Palius & Maher, 2013). Although results from these studies using video-based assessment data are not conclusive, they offer evidence of promise of the VMC videos as resources for learning about students' mathematical reasoning. These findings have led to shifting our focus to the prospects of the VMCAnalytic tool for assessment.

Technology-Enhanced Assessment

Learning technologies are providing new opportunities to teach thinking and reasoning and what students can be expected to do to show their knowledge and skills (Pellegrino & Quellmalz, 2010). These changes allow us to think about what is assessed and new ways to provide evidence of understanding. In particular, technology allows assessing a range of complex performances (Pellegrino, 2013) An important source of evidence for assessment can be found in the artifacts that learners create with technology (de Jong, Wilhelm, & Anjewierden, 2012). In particular, technology provides a high level of expressiveness as learners can create multimedia artifacts. There is also evidence that assessment based on teacher's analysis of video can predict student learning outcomes. For example, Kersting, and colleagues (Kersting, Givvin, Sotelo, & Stigler, 2010) used a video analysis task to assess teacher knowledge. To score the teacher's analyses, they used a rubric that measured Mathematical Content, Student Thinking, Suggestions for Improvement, and Depth of Interpretation on a 3-point scale. They found that suggestions for improvement predicted student learning. Unlike the Kersting et al. study, in our work, the VMCAnalytic is an embedded assessment that is both designed to support learning and to provide an occasion for assessment.

The VMCAnalytic provides opportunity for students to make their thinking visible and provides unique opportunities for assessment because students must bring together conceptual knowledge and the rich videos of learning in action (Derry, Hmelo-Silver, Nagarajan, Chernobilsky, & Beitzel, 2006). As Pellegrino & Quellmalz (2010) note, technology offers opportunities for innovative assessment of complex skills along while allowing scaffolding that promotes learning. Although the VMCAnalytic provides opportunities for eliciting complex performance, we consider this to be part of an instructional system that includes teacher scaffolding and rubrics that make expectations clear for both learners and instructors. The VMCAnalytic thus provides a means for instructors to monitor what students are learning and help scaffold their progress as a formative assessment and can provide evidence of a student's (developing) competence. Although space precludes providing examples of VMCAnalytics here, we refer the reader to examples of published VMCAnalytics: http://bit.ly/1hZcjoR.

Methods

The data analyzed in this paper come from 63 VMCAnalytics that were created over the last two years by participants in 7 courses. The VMCAnalytics were graded on an integer scale from 0 to 3 on two levels; a local individual event level and a global level evaluating the VMCAnalytic as a whole. Each event that contributed to the participants VMCAnalytic was rated on how well the event fit into the overall description. Examples of high scoring events included text in the description that explained how the video they chose lent support to their overall description. Lower scoring events tended to select video but not situate it or made faulty inferences from the video. The scoring rubric is shown in Table 1. Two independent coders with expertise in evaluating and creating VMCAnalytics scored the 63 VMCAnalytics. Inter-rater reliability between the two coders was 88.72%.

Table 1. Scoring rubric

Criteria	0	1	2	3
Overall description of the analytic is very explicit about what it shows	Off topic	On topic, but vague about what analytic will show	Discusses topic explicitly, but is only related peripherally to topic and does not capture its essence	Captures essence
Each event contributes meaningfully to the overall purpose of the analytic	Most events are extraneous or weakly linked to purpose	Most events contribute well, some extraneous and have no connection	Most events contribute well, some are weakly linked	All events contribute strongly to the purpose
Clips connect to each other in a meaningful way	No easily discernible logical sequence	Some events are in sequence, most are not	Most events are connected, some seem unconnected	All events are in logical sequence
Claims are backed with evidence	No claims made are backed by evidence	Some claims backed by evidence but most are not	Most clams backed by evidence, a few are not	All claims made in descriptions backed by video evidence or research literature
Overall clarity of analytic	Descriptions are all hard to understand/ unclear	Some descriptions easy to understand, most are difficult	Most descriptions easy to understand, some difficult, or overall description unclear	Easy to understand the intent of each description as well as the intent of the overall description
Overall coherence	Hard to understand why any of the events were included or how they contribute to purpose	Hard to understand why most of the events are included but some are easy	Easy to understand why most of the events are included but some are hard	Easy to understand why each event included. Overall description describes purpose well.
Mathematical/ Learning Sciences depth	Does not address learning	Superficial use of terminology	Mid-level	Builds on specific learning theory
Fit of Title	Off topic	Vague, cannot predict content based on title	Related to topic peripherally, does not capture its essence	Captures its essence

Because the score a participant received may have been related to the context of the assignment, below we describe the classes on which the data were collected. VMCAnalytics were collected from three semesters of Introduction to Mathematics Education, which is a required class for students to obtain an M.Ed, Ed.D., or Ph.D. degree with a specialization in mathematics education. Several participants were in other degree programs taking the class as an elective. The course used a hybrid format with a mix of in-person meetings and online asynchronous discussions. During in-person sessions, participants worked in groups on problem-solving tasks and shared solutions. For homework, they watched videos on the VMC of students working on the same or similar tasks, read related articles from research literature, and engaged in small-group discussions online

The means by which the VMCAnalytic project in Introduction to Mathematics Education was presented to the students varied for each semester. The participants in the first iteration of the course (n=11) worked initially with a subset of VMC videos about the Guess My Towers task (see videomosaic.org for all video examples). In this task, students must determine and compare various probabilities of different events occurring when building towers four-tall using Unifix cubes while selecting from two colors. The VMC has a series of five videos with fifth graders working on the task over an eighty-minute session. The participants in this iteration of the class worked in pairs creating VMCAnalytics around this task with the constraint that they have a minimum of three events and a maximum of six. The run time of their VMCAnalytic was also to be between four and ten minutes. In the second iteration of the class (n=19), the participants were free to create a VMCAnalytic about any topic. Individuals created their own VMCAnalytic, but groups were formed along common themes. Several participants were interested in constructing VMCAnalytics for professional development around building fraction concepts and met several times throughout the semester for sharing their VMCAnalytics with their group and a course instructor. The group discussions were centered on helping each other improve their individual VMCAnalytics. Similar groups were also formed for teaching algebra concepts and constructing examples with events that demonstrate how to implement the Common Core Standards for Mathematics. In the third iteration during the spring of 2013 (n=11), students were free to use any videos they wanted but were constrained to a ten-minute run time on their VMCAnalytic.

The Critical Thinking and Reasoning course (n=8) is an online, mathematics education elective for graduate students. In this class students watched videos of a group of fourth graders over several months as they explored fraction ideas before they were introduced formally (Palius & Maher, 2013) The participants drew from these videos to construct their VMCAnalytics.

The Early Algebraic Learning course (n=7) is a graduate mathematics education elective. Videos of students engaging in the Guess My Rule task were used extensively in the course. Many participants drew on those and related videos to construct their VMCAnalytics. As in the spring sections of Introduction to Mathematics Education, the participants in this course had a time limit of ten minutes imposed for their VMCAnalytics.

The Design-based Research (DBR) course (n=7 across two semesters) used the VMCAnalytic as part of a brief exercise in video analysis. This course consisted of doctoral students in a range of disciplines and who were focused on research methods rather than mathematics education. We expected these students to focus on more general aspects of learning and collaborative knowledge construction. This group of students was expected to provide a contrast with the other classes. In the 2011 class, limited directions were provided. This was addressed in the subsequent iteration when the assignment was more structured and some students worked in a group. Students were pointed to a limited set of videos rather than the whole repository and were given greater directions as to number of events, length, and the need to make connections to learning theories.

Results

Table 2 contains mean scores and standard deviations for each metric across the different classes. Pairwise differences were computed using the Wilcoxon signed-rank test and found significant differences (all $p < 0.05$) between DBR Fall 2011 and all the other classes except Critical Thinking and Reasoning across overall description, the clips connecting meaningfully, claims are backed, overall clarity and coherence, and event relevance, suggesting that the students with the least preparation and direction had the lowest scores. Significant differences were also found in the relevance of the events between Introduction to Mathematics Education in Spring 2012, where a smaller subset of videos were used, compared with all of the other classes except Introduction to Math Education Fall 2012. This suggests that using the smaller number of videos was beneficial. We conjecture that this is because it reduced the amount of search in which the course participants needed to engage. Moreover, the Introduction to Math Education Fall 2012 class, which was involved in developing the rubrics, did better than all the other classes besides Spring 2012 on events connect meaningfully, claims backed with evidence, and overall clarity/coherence. This suggests that the use of the rubric helped focus student efforts. Although the numbers are too small for a statistical comparison, inspection of the scores for DBR 2012 compared with the 2011 course suggests that the increased structuring of the assignment led to higher quality artifacts being produced.

Table 2. Mean Ratings of VMCAnalytics across classes (standard deviations in parentheses)

Class	n	Overall description	Events connect meaningfully	Claims backed w/ evidence	Overall clarity and coherence	Math depth	Learning Science depth	# of events	Event relevance average
Early Algebraic Learning	7	2.57 (0.53)	2.29 (0.48)	2.14 (0.69)	2.14 (0.38)	1.86 (0.69)	1.71 (0.76)	6.86 (2.03)	2.06 (0.50)
Design Based Research Fall 2011	3	1.00 (0)	1.00 (0)	1.00 (0)	1.99 (0)	1.00 (0)	1.00 (0)	3.00 (3.00)	0.78 (0.69)
Design Based Research Fall 2012	4	2.00 (0.82)	2.00 (0.82)	2.00 (1.15)	1.75 (0.96)	1.75 (0.5)	2.25 (0.50)	7.00 (2.45)	1.94 (0.92)
Reasoning and Critical Thinking	8	2.00 (0.76)	2.06 (0.68)	2.13 (0.64)	2.13 (0.64)	2.00 (0.76)	2.00 (0.53)	9.38 (5.78)	2.27 (0.55)
Introduction to Mathematics Education Spring 2012	11	2.55 (0.82)	2.64 (0.67)	2.46 (0.69)	2.73 (0.61)	2.19 (0.61)	1.64 (0.50)	4.09 (0.70)	2.71 (0.86)
Introduction to Mathematics Education Fall 2012	19	2.52 (0.61)	2.42 (0.61)	2.47 (0.70)	2.42 (0.61)	2.21 (0.71)	1.95 (0.78)	7.26 (2.58)	2.35 (0.53)
Introduction to Mathematics Education Spring 2013	11	2.46 (0.52)	2.00 (0.89)	2.00 (0.89)	1.91 (0.83)	2.37 (0.81)	1.27 (0.65)	8.27 (2.69)	2.12 (0.63)

Table 3. Correlation between measures

	Overall description	Events connect meaningfully	Claims backed w/evidence	Overall clarity /coherence	Math depth	Learning Sciences depth	Number of events	Event relevance average
Overall description	1	0.60	0.72	0.66	0.47	0.38	0.13	0.67
The events connect meaningfully	0.60	1	0.77	0.91	0.56	0.34	-0.09	0.86
Claims are backed with evidence	0.72	0.77	1	0.8	0.48	0.47	0.01	0.75
Overall clarity and coherence	0.66	0.91	0.8	1	0.55	0.33	-0.08	0.87
Mathematical depth	0.47	0.56	0.48	0.55	1	0.092	0.25	0.61
Learning Sciences depth	0.38	0.34	0.47	0.33	0.092	1	-0.05	0.30
Number of events	0.13	-0.09	0.01	-0.08	0.25	-0.05	1	-0.05
Event relevance average	0.67	0.86	0.75	0.87	0.61	0.30	-0.05	1

Correlations were calculated across rating metrics (Table 3). High correlations were found between the clips connecting meaningfully and overall clarity and coherence (r=0.91). This suggests that students who make meaningful connection across events are also more likely to have an overall coherent VMCAnalytic. Similarly, the correlations between clips connecting meaningfully and event relevance (r=0.86), and overall clarity and event relevance (r=0.87) suggests that selecting relevant events is another important factor in the overall clarity of VMCAnalytics. There is a low correlation between the number of events and all of the other metrics suggesting that they are measuring different aspects of learner performance.

Discussion

The VMCAnalytic shows promise of being a useful tool in a system of formative and summative assessment. Constructing multimedia artifacts was an integral component of the instructional design for each of the courses in which they were used, making learner thinking visible and open for discussion and revision. However, equally important is what this research reveals about the kinds of structures and scaffolds that the use of the VMCAnalytic can provide. The rubric provides clear expectations and a roadmap for student use in creating this multimedia artifact as our results from Introduction to Mathematics Education Fall 2012 suggest. Moreover, structuring the task by reducing the amount of video that students need to search also appears to be beneficial as demonstrated by both the Introduction to Mathematics Education Context and Design-Based Research results. For instructors, students' evolving understanding becomes transparent and provides new insights into a student's intellectual journey in thinking critically about children's mathematical thinking and reasoning.

We are now studying the advantages of the opportunities that this assessment provides. As a follow up to this research, a summer research practicum course served as a context for further refinement of the multimedia artifacts made by a subset of participants from the previous semester's Introduction to Mathematics Education and Early Algebraic Reasoning courses. One of the researchers who scored the VMCAnalytics met with graduate students to discuss how their VMCAnalytics were scored using the rubric, and a Senior member of the research team worked with those students to refine their analytics by working on them as a collaborative group focusing on one artifact at a time. An online forum supported the summer practicum as students gave each other feedback based on rubric criteria. These learners shared their work with other members of the practicum community (who worked on different projects) midway and at the end of the term. The next phase of our research will entail detailed analysis of the revision process and how it might have been scaffolded by the rubric as well as current tools and forthcoming technology affordances being designed to support such work.

In our ongoing research, we are collecting additional data to provide process feedback to instructors on how the students are using the VMC as they create VMCAnalytics. Such information offers promise to provide automated analyses to instructors that would support targeted facilitation of student learning. These analyses will draw from log data of student search, iterative refinement of VMCAnalytics, and other forms of learning analytics. In addition, collaboration tools such as threaded discussions and blogs will add opportunities for peer assessment and student reflection.

As we have shown here, the VMCAnalytic is part of a system of learning and assessment. It affords opportunities for learners to make their thinking visible and available for discussion, refinement, assessment, and revision (Collins, 2006). The rubrics can be helpful for guiding students as to what the expectations as well as for researchers in their evaluations. These rubrics help provide constraints on the task that channel the learners in productive ways (Reiser, 2004). Our results demonstrate that providing guidelines for the assignment to create VMCAnalytics are important. As we continue developing and refining the tool as well its use in assessment, we are enthusiastic about the promise of the VMCAnalytic tool.

References

Agnew, G., Mills, C. M., & Maher, C. A. (2010). *VMCAnalytic: Developing a collaborative video analysis tools for education faculty and practicing educators.* Paper presented at the 43rd Annual Hawaii International Conference on System Sciences, Kauai, HI.

Bielaczyc, K. (in press). Informing design research: Learning from teacher's designs of social infrastructures. *Journal of the Learning Sciences*. doi: 10.1080/10508406.2012.691925

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education., 24*, 417-436.

Bransford, J. D., Derry, S. J., Berliner, D., Hamerness, K., & Darling-Hammond, L. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. D. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 40-87). San Francisco: Jossey-Bass.

Collins, A. (2006). Cognitive Apprenticeship. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 47-60). New York: Cambridge University Press.

Collins, A., & Halverson, R. (2009). *Rethinking Education in the Age of Technology: The Digital Revolution and Schooling in America*. New York: Teachers College Press.

de Jong, T., Wilhelm, P., & Anjewierden, A. (2012). Inquiry and assessment; future developments from a technological perspective. In M. Mayrath, D. Robinson & J. Clarke-Midura (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern researc* (pp. 249-265). Charlotte, NC: Information Age Publishing.

Derry, S. J., Hmelo-Silver, C. E., Nagarajan, A., Chernobilsky, E., & Beitzel, B. (2006). Cognitive transfer revisited: Can we exploit new media to solve old problems on a large scale? *Journal of Educational Computing Research, 35*, 145-162.

Hmelo-Silver, C. E., Maher, C. A., Palius, M. F., Sigley, R., Alston, A., Agnew, G., & Mills, C. (2013). Building multimedia artifacts using a cyber-enabled video repository: The VMCAnalytic. *46th Hawaii International Conference on System Sciences* (pp. pp. 3078-3087). Hawaii: IEEE.

Kafai, Y. B., & Ching, C. C. (2001). Affordances of Collaborative Software Design Planning for Elementary Students' Science Talk. *Journal of the Learning Sciences, 10*, 323-363.

Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers analysis of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education, 61*, 172-181.

Maher, C. A., Powell, A. B. & Uptegrove, E. (2010). Combinatorics and reasoning: Representing, justifying and building isomorphisms. New York: Springer Publishers.

Maher, C. A. (2008). Video recordings as pedagogical tools in mathematics teacher education. In D. Tirosh and T. Wood (Eds.), *International Handbook of Mathematics Teacher Education: Vol. 2: Tools and Processes in Mathematics Teacher Education (pp. 65-83)*. Rotterdam, The Netherlands: Sense Publishers.

Maher, C. A., & Davis, R. B. (1995). Children's explorations leading to proof. In C. Hoyles & L. Healy (Eds.), *Justifying and proving in school mathematics* (pp. 87-105). London: Mathematical Sciences Group, Institute of Education, University of London.

Maher, C. A., & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education, 27* 194-214.

Maher, C. A., Palius, M. F., Maher, J. A., Hmelo-Silver, C. E., & Sigley, R. (2014). Teachers can learn to attend to students' reasoning using videos as a tool. *Issues in Teacher Education 23* (1), 31-47.

Maher, C. A., Powell, A. B., & Uptegrove. (2010). *Combinatorics and reasoning: Representing, justifying and building isomorphisms*. New York: Springer Publishers.

Palius, M. F. (2013). *Deepening teachers' awareness of students' mathematical reasoning through video study in an online course. Rutgers University.* Unpublished doctoral dissertation, Rutgers University.

Palius, M. F., & Maher, C. A. (2013). Teachers learning about student reasoning through video study. *Mediterranean Journal of Research in Mathematics Education, 12*, 39-55.

Pellegrino, J. W. (2013). Measuring what matters: Technology and the design of assessments that support learning. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood & N. Winters (Eds.), Handbook of Design in Educational Technology (pp. 377-387). New York: Routledge.

Pellegrino, J. W., & Quellmalz, E. S. (2010). Perspectives on the integration of technology and assessment. *Journal of Research on Technology in Education, 43*, 119-134.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences, 13*, 273-304.

Rodriguez, P. M., Frey, C., Dawson, K., Liu, F., & Ritzhaupt, A. (2012). Examing student digital artfacts during a year-long technology integration initiative. *Computers in the Schools, 29*, 355-374.

Schwartz, D. L., & Hartman, K. (2007). It's not television anymore: Designing digital video for learning and assessment. In R. Goldman, R. D. Pea, B. J. S. Barron & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 335-348). Mahwah NJ: Erlbaum.

Yackel, E. & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, G. W. Martin, and D. Schifter, (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 227-236). Reston, VA: National Council of Teachers of Mathematics.

Zahn, C., Pea, R., Hesse, F. W., & Rosen, J. (2010). Comparing simple and advanced video tools as supports for complex collaborative design tasks. *Journal of the Learning Sciences, 19*, 403-440.

Zhang, M., Lundeberg, M., & Eberhardt, J. (2011). Strategic facilitation of problem-based discussion for teacher professional development. *Journal of the Learning Sciences, 20*, 342-394. doi: 10.1080/10508406.2011.553258

Acknowledgments

The Video Mosaic Collaborative is a research and development project sponsored by the National Science Foundation grants DRL 0822204 and 1217087). We gratefully acknowledge the NSF and note that the views expressed here are those of the authors are not necessarily those of the NSF.

Varied Appropriations of Tools from Professional Development: Moving Beyond Levels

Huy Q. Chung and Elizabeth A. van Es, University of California, Irvine, CA, USA
Email: hqchung@uci.edu, evanes@uci.edu

Abstract: In this empirical study we trace teacher's appropriation of literacy teaching tools from professional development and the varied ways they enacted these tools in their classrooms. Teacher appropriation is important in studying professional development because far too often studies do not account for teacher learning and how this learning sustains reform-minded practices. Using two years of qualitative data we traced the appropriation of tools and their impact on instruction. Pushing on Grossman, Smagorinsky, and Valencia's (1999) five levels of appropriation, results indicate that the teachers appropriated tools for different pedagogical reasons, such as appropriating the tools to organize their students' learning, to accommodate their students' learning needs, and to integrate other professional development initiatives. Implications on designing and sustaining teacher learning in professional development are discussed within.

Introduction

In the past decade, researchers, policymakers, and practitioners have given increased attention to the impact of teachers on student achievement (Baker et al., 2010; Darling-Hammond et al., 2009; Yoon et al., 2007). The logic behind this focus stems from studies that conclude that teachers significantly affect student achievement (Nye, Konstantopoulos, & Hedges, 2004). Teacher effects matter even more for schools in lower socio-economic communities that experience high teacher attrition. In order for systematic change to happen in schools and to sustain a strong culture of teaching, education researchers call for an investment in teachers and their practice (Ball & Forzani, 2011; Morris & Hiebert, 2011). Large amounts of funding have been provided to schools and districts to offer professional development programs to improve teaching in hopes of increasing student learning (von Zastrow, 2010). A substantial amount of research related to these efforts focuses on the impact of professional development on student achievement (Baker et al., 2010; Guskey, 2002; Wallace, 2009; Yoon et al., 2007). This emphasis is clearly important; policymakers and taxpayers deserve to know if such efforts are making a difference for students. However, we argue that ensuring lasting impact of teachers' practice and long-term improvements on student learning requires a comprehensive understanding of, and attention to, teacher learning and the relationship between teachers' long-term learning and classroom practices.

While research on teacher education points to the complexity and longevity of developing teacher expertise, studies of the effectiveness of professional development on student achievement have ignored this critical piece of the puzzle. To be clear, research in this vein does look at factors that contribute to teacher learning, such as time spent in professional development, opportunities for active learning and collaboration, and instructional support (Desimone et al., 2002; Garet et al., 2001; Guskey, 2002). However, these studies do not examine how these settings provide or limit teachers' opportunities to learn, the nature and development of that learning, and how the contexts in which they work interact with their participation and learning in professional development. We argue that research that examines the impact of professional development not only consider the effect on student performance, but also what and how teachers learn in professional development contexts, and how that impacts their thinking and practice, which is consistent with other research that advocates for a more comprehensive approach to studying the effects of professional development on teachers (Desimone, 2009; Opfer & Pedder, 2011). In this study, we propose a model to broaden the conception of impact studies to account for several components that influence both what and how teachers learn and develop in their profession. Specific attention to teachers' appropriation of pedagogical tools will be used as an example of this broadening. If teachers are so essential to student learning as research suggests, then attention to teachers' learning, development, and well-being are equally important to consider.

Literature Review

The last decade of research on teacher effectiveness has provided important insight into factors that have the potential to improve teaching and in turn impact student achievement, such as opportunities to collaborate and a focus on student learning outcomes (Guskey, 2002; Thompson, Windschitl, & Braaten, 2012). Three of the main contributions from this literature are the knowledge that most teachers experience professional development on a regular basis, but what that looks like varies across school contexts (Garet et al., 2001; Vescio, Ross, & Adams, 2008), the establishment of an empirical base for what makes for effective professional development (Darling-Hammond et al., 2009; Desimone et al., 2002; Garet et al., 2001), and the conclusion that professional development can impact student achievement (Darling-Hammond, 2000; Guskey, 2002; Wallace, 2009).

Despite these contributions, these studies are also limiting in that they do not capture what and how teachers learn, or the sort of useable knowledge they develop for their teaching. They also neglect how the local, school, and district contexts influenced teachers' ability to take up the ideas they learned in professional development for their practice over time. Teaching is a complex activity. It requires an extensive knowledge-base about teaching, content, learners, and school contexts (Shulman, 1986), development of beliefs and identities as teachers (Richardson, 1996), and constant improvement of teaching practices through critical analysis and reflection. As Schoenfeld (2011) argues, learning to teach, and developing expertise in teaching takes time and is slow to develop.

Impact studies fail to take the complexity and time investment of learning to teach into account and thus often find selected outcomes (i.e., student achievement) plateau or fade after the initial year(s) of professional development (Borman, 2005). Moreover, institutional memory and the expectations that arise out of the cultures of schools also influence the impact of professional development on teachers' practice (Grossman, Smagorinsky, & Valencia, 1999). As with any reform, policy climates change, schools respond differently, and interpret initiatives to meet their context. Thus, another reason why meaningful learning gains may not be sustained is that the designs do not account for the complexity of schools responding to policy climates, the challenges of building a climate of learning and improvement for all students, and school leaders guiding teachers in focused improvement on teaching practice (McLaughlin & Talbert, 2006).

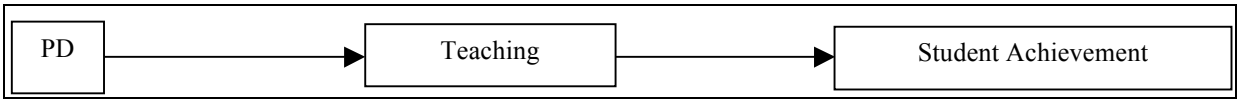


Figure 1. Limited Conception of Impact of Professional Development

In sum, research on the effectiveness of professional development on student achievement attempt to streamline the impact process, but do not currently account for the complexity of teacher learning and changing teaching practices (Opfer & Pedder, 2011). Figure 1 illustrates this limited conceptual model for studying the impact of professional development. In this model, the teacher is largely absent. Instead, the focus is on student outcomes, and teachers' practice (teaching) is used to explain these results. To put it another way, the professional development program is designed to leverage certain teaching practices that have been identified to improve student achievement. Often times, these practices are advocated for without attention to the context of where these practices will be enacted or by whom. If student achievement improves then the professional development program is seen as successful; however, these successes may ignore other important factors that account for their success such as contextual factors, particularly teachers, their experiences, and learning.

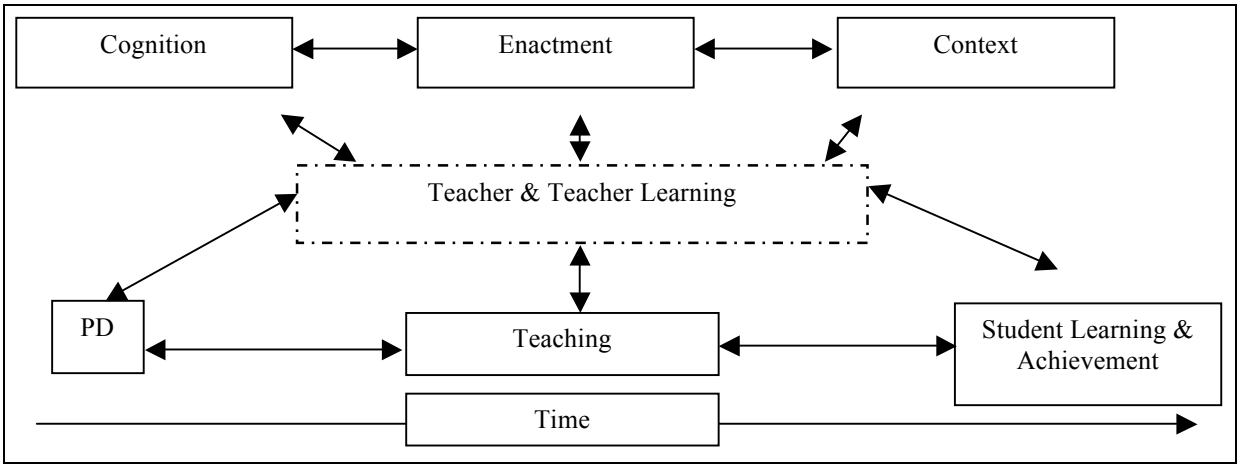


Figure 2. Expanded Conception of Teacher Development

Importantly, across much of this research, studies of teacher learning are largely absent and little research examines how teachers were supported while they implemented new teaching practices (Penuel, Gallagher, & Moorthy, 2011). The model we propose (see Figure 2) incorporates many of the factors we described that operate to influence teacher learning. These factors include:

- *Teachers' cognition*, including knowledge, beliefs, and identity;
- Teachers' local *enactment* of practices promoted in professional development;

- The local, state, and national policy *context*, as well as the school culture and community on teachers' learning;
- *Time* invested for sustained change in teaching and student learning

In this model, the link between professional development, teaching and student achievement remains. However, additional elements are incorporated to expand the conceptualization of how researchers and policy makers should seek to understand the impact of professional development on teachers. Moreover, in this model, we highlight the reciprocity between these elements. We ground our argument using research on how learning occurs in the interactions that take place among multiple reference points. These reference points include teacher cognition, teacher enactment, and teacher context which are understood as developing and interacting with each other over time. The following paragraphs provide a brief, but concise review of why these factors matter for research on teachers participating in professional development, with specific attention to teachers' appropriation of pedagogical tools.

Research in teacher education shows that knowledge for teaching matters and plays a significant role in influencing student learning (Hill et al., 2008). Shulman's (1987) seminal work identifies at least seven different types of knowledge that teachers need to develop to be effective and highlight how they are inter-related and difficult to develop. Teacher beliefs are another aspect of teacher cognition that influences teaching (Pajares, 1992; Richardson, 1996; Stipek et al., 2001). Richardson (1996) underscores the fact that teacher beliefs are strong and enduring and develop from different life experiences that can either be personal, experienced during schooling, or developed in formalized settings such as in their teacher education programs and drives behavior.

In regards to teacher enactment, we also adopt Shulman's (1987) vision of teaching and teacher education, arguing that teachers are professionals who are capable of "enacting –of acting in a manner that is self-conscious with respect to what their act is…" (p. 13). In other words, teachers do not just do what others tell them to do nor do they simply adopt the strategies they learned as students; rather, they are aware of their thinking as they teach and they make conscious decisions about their practice based on what they know and believe (Schoenfeld, 2011). Consistent with Shulman (1987) and the situative perspective (Wenger, 2010), the process in which this change happens can be characterized as an appropriation for practice. Appropriation (Rogoff, 1995) refers "to the process by which individuals transform their understanding of and responsibility for activities through their own participation" (p. 147) in group settings. The interactions that occur during these joint activities shape both the group and community as a whole, as well as the individual. This individual shaping is the appropriation process. During this process individuals discuss and come to agree on shared understandings around particular artifacts, processes, and language and take what is most useful for them during an activity to serve their own purposes. Their knowledge and beliefs also shape what they come to appropriate (Rogoff, 1990). In the context of professional development, even though the group may have an agreement, each individual will come to appropriate tools from professional development differently (Grossman, Smagorinsky, & Valencia, 1999). In other words, teachers will modify and adapt what they learn in professional development based on their knowledge, beliefs, and context (Grossman, Smagorinsky, & Valencia, 1999; Rogoff, 1995). In some cases, the intent remains (Neuman & Cunningham, 2009) and in other cases, teachers greatly modify the information learned and their interpretation of the practice is not in the spirit of the original professional development program (Brown & Campione, 1996).

Related to teacher enactment is the idea of tools assisting teachers in their teaching. Situated learning theory explains that tools embody knowledge valued by their creators and community of users (Pea, 1993). However, individuals take them up and use them in ways that are most suited to their needs. Curriculum materials are one example. Remillard's (2005) review of curriculum use explains that historically, curriculum materials have been viewed as a means to reform teaching practice. However, research on curriculum use reveals mixed findings. Some research suggests that teachers embrace new materials, while others reject and subvert the goals of curriculum. Remillard (2005) proposes that this is the case because the "teacher-curriculum relationship is intertwined with other teacher practices, is dependent on the particular teacher and curriculum, and is situated in a specified context" (p. 212). Attending to this relationship is necessary for understanding the success of materials to transform teaching.

Research on teacher cognition and teacher enactment draws attention to the individual teacher. A third aspect we include in the model is the context in which the teacher is situated. Here, we draw attention to the influence of the professional development contexts, the organizational contexts, and the policy contexts on teacher learning and changes in practice (Talbert & McLaughlin, 1994). Situated learning theory posits that teachers are nested in contexts and within each context they must learn to negotiate the "rules" for each. Each context can be considered as a system with inherent goals. However, actors within these systems also must work together to create mutually negotiated practices that are within cultural norms and boundaries, are agreed upon, and concrete.

One of the criticisms of current studies of the impact of professional development on students is that they are limited in timeframe (Baker et al., 2010). That is, these studies look at the year that teachers are in

professional development and then look at test scores or teacher change in practice that same year. The presumption behind these studies is that professional development will impact student learning in that same year. However, Richardson and Placier (2001) point out that teacher change is neither linear nor predictable but is rather more idiosyncratic and can take years to manifest in practice. Webster-Wright (2009) also argues that professional development studies need to consider what it takes for authentic teacher learning, citing that time is an essential need and that teacher learning is an indefinite process. Darling-Hammond et al. (2009), Garet et al. (2001), Guskey (2002), and Hawley and Valli (1999) identified the need for sustained and supported professional development opportunities that are beyond one-shot workshops (Goldenberg & Gallimore, 1991). Thus, we argue that studies of the effectiveness of professional development on student achievement and learning take into account time for teacher learning that includes teacher cognition, enactment, and contexts.

Situated learning theory suggests that learning is not unidirectional. Instead, the communities in which people participate both shape and are shaped by the members of the group (Brown, Collins, & Duguid, 1989). Additionally, Kazemi and Hubbard (2008) argue that research attend to the co-evolution of teachers' participation across contexts. That is, they contend that teachers will take up what they learn in professional development in different ways and thus changes in practice will vary across participants. Because of this variation, they will bring back to subsequent professional development activities different experiences, which will influence how and what they learn in future professional development activities. The model we propose gives attention to this movement because it is critical for understanding if and how teachers are impacted by professional development and thus improve student learning.

Finally, this model adds student learning as an additional outcome of interest. While student achievement is critical, current approaches to measuring achievement are narrow in focus (Baker et al., 2010). Additionally, current conceptions of learning suggest that learning is not just about pieces of knowledge, but also includes how learners manage their learning and knowing, how they develop practices for participating in a community, and how they learn to use tools and resources for productive participation (Brown, Collins, & Duguid, 1989). These dimensions are needed in studies of the impact of professional development on teachers and students because without acknowledging these dimensions as critical components, professional development initiatives have little chance of having a lasting influence (Borko, 2004; Desimone, 2009).

Study Context

For this qualitative study, we utilize our expanded conception of teacher development (see Figure 2, above) to frame our study on the impact of a two year professional development program, the Pathway Project, on teachers' appropriation of tools, namely, cognitive strategies for reading and writing analytically. Cognitive strategies are conceptual tools and processes that can help students become more meta-cognitive about their work (Olson & Land, 2007). In her book, *The Reading/Writing Connection*, Olson (2011) likens the cognitive strategies as tools within a tool kit where they are drawn upon depending on the task at hand. Some cognitive strategies introduced in the Pathway Project include planning and goal setting, tapping prior knowledge, revising meaning, adopting an alignment, and making connections.

These strategies are disseminated through teacher professional development. The intent of the program is to provide teachers with lessons and materials to introduce the cognitive strategies to students towards the intended goal of writing analytical essays around either fiction or non-fiction texts, collaboration and support in the use of cognitive strategies, and additional literacy professional development such as strategies on how to teach poetry, use Socratic seminars, or developing academic vocabulary. This professional development program is the perfect context for us to study the following question: How do teachers appropriate and enact the Pathway Project tools for use in their teaching of writing over time?

The teachers in this study come from the two lowest-performing middle schools from an urban school district serving a majority population of under-represented minorities (98% Latino; 88% low-income; 80% English language learners [ELL]; and 74% qualify for free and reduced price lunches [FRPL]). The two middle schools are Lion Middle School (ELL: 59%; FRPL: 92%) and Sparrow Middle School (ELL: 56%; FRPL: 92%). The entire English language arts department from both schools participated in three professional development contexts for two years, during Academic Years 2011-2012 and 2012-2013. Each year, 32 teachers attended a series of 6 full release days, a series of 5 two-hour after school workshops, and 36 weekly grade-level meetings involving on-going professional development around the Common Core State Standards, cognitive strategies use, and instruction around analytical writing. Of note is the fact that a total of 48 unique teachers participated in the Pathway Project, but only 17 teachers were able to participate in the Pathway Project for two years with 15 teachers laid-off the first year of participation and 15 new teachers hired during the second year of participation. Data will only come from teachers who participated both years.

Data Sources

Approaching this question requires comprehensive qualitative data analysis using a variety of sources of data to understand how the teachers' understood the intent behind the Pathway Project tools as well as how they take

them up and use them in their teaching. The data sources for this study consisted of two years of field notes of the professional development settings, 80 observational protocols from literacy coaches, 34 teacher self-reports during written reflections, and 6 teacher focus group interviews. The field notes came from 12 full release days and a series of 10 two-hour after school workshops over two years. The protocols were conducted on random announced days by two literacy coaches. Their feedback came in the form of letters that outlined what learning activities took place during instruction and how the teachers incorporated or did not incorporate the Pathway Project tools while teaching. The teacher self-reports were reflections on the Pathway Project, what tools they used from the Pathway Project, which tools they found the most valuable, and how their teaching has changed as a result of the Pathway Project. Finally, the focus group interviews were conducted at the end of each year of implementation. They were done in grade level groups and addressed the affordances and limitations of the Pathway Project on the teachers' instruction.

Analytic Methods

Analysis was done in three phases. During the first phase a comprehensive list of tools, lessons, and materials provided to teachers during professional development meetings was generated through field notes, the observational protocols, and questions from teacher reflections practice surveys. The three types of tools identified in this phase include tangible tools, conceptual tools, and experiential tools (e.g., unit and lesson plans). During the second phase, the first author coded the 80 feedback letters from the literacy coaches for all 17 teachers and traced the presence of the tools identified during the first phase and noted how they were being used. The first author also used the surveys and the focus group interviews collected during both years for all teachers to triangulate the data. In the third phase, the first author qualitatively analyzed how these tools were appropriated, adapted, or adopted by comparing their enactment to field notes of how these tools were meant to be used when introduced during the professional development days. After analysis, the second author pushed for confirming and disconfirming evidence as well as for robust examples of these appropriations to develop validity across the data set.

We approached the data with a top down, bottom up approach and analyzing them for concepts (Corbin & Strauss, 2008) by considering how a wide range of tools are used or not used, when they are used, and to what extent they were appropriated from the Pathway Project to understand the degree in which the principles behind these tools have been internalized when in practice (Rogoff, 1990). Moreover, we also characterized how the teachers utilized these tools in different ways by triangulating their use with other data and constantly comparing their use (Miles & Huberman, 1994) with, for example, the focus group interviews or teacher reflection surveys. In this study we also move beyond the five levels of appropriation that Grossman, Smagorinsky, and Valencia (1999) proposed as a framework to study teachers' appropriation of tools. Though the levels are a good guide, we argue that a more comprehensive understanding of how teachers appropriate tools and for what purposes is more useful in designing professional development for teachers.

Results

In general, there were five distinct patterns in which the teachers appropriated the Pathway Project tools. The five patterns are grouped around the purposes for appropriation: appropriation as organizing student learning, appropriation as incorporating competing professional development goals, appropriation as accommodation of student learning needs, appropriation as addressing gaps in the curriculum, and appropriation as engaging students. The following paragraphs provide examples of how teachers appropriated the Pathway Project tools in these ways, demonstrating the complexity of enacting professional development.

Appropriation as Organizing Student Learning

The majority of appropriations came in the form of teachers creating different tools pulling disparate components of the Pathway Project tools together into packets that walked students through the reading, interpreting, and writing of complex textual analysis emphasizing the use of cognitive strategies, graphic organizers, writing prompts, and essay components. Other ways they organized their students' learning was creating a check list of tasks, rearranging the sequence of the curriculum, or only using the basic components of a curriculum to address the demands of time and pacing.

Appropriation as Incorporating Competing Professional Development Goals

While the Pathway Project tool was an important focus for the teachers in this study, their district also had other professional goals for the teachers. In order to address the demands of both the Pathway Project and their other professional development activities, some teachers would combine different tools together to meet these demands. For example, thinking maps and graphic organizers, a district-wide initiative, were incorporated into the texts that the teachers were given from the Pathway Project. Other teachers would also use the cognitive strategies with Sheltered Instruction Observation Protocols (Echevarria, Short, & Powers, 2006).

Appropriation as Accommodation of Student Learning Needs

Another prevalent pattern of appropriation was adjusting the Pathway Project tools as accommodating student learning needs. A common theme from the focus group interviews revealed that the teachers felt the tools from the Pathway Project were too advanced for the majority of their students who were either English language learners, special-needs students, or below grade level in terms of reading and writing. They would often change the complexity of the curriculum materials by reducing the cognitive demands of writing prompts, by replacing texts given to them with other texts with less demanding language, by chunking lessons across multiple days rather than one single day as they were intended to be given, and by creating new tools with more scaffolds than were given to them.

Appropriation as Addressing Gaps in the Curriculum

Some teachers also addressed gaps in the curriculum given to the teachers. The curriculum from the Pathway Project addressed the teaching of analytical writing, specifically through the analysis of theme in both fiction and non-fiction texts. The curriculum also included ways to scaffold students in the different components of an analytical essay, such as an introduction, body paragraphs, conclusion, a thesis about the theme, and textual evidence. These lessons all used different texts to be taught over two or three days. Some of the teachers felt that the curriculum was not comprehensive enough and would develop other tools or materials to address such gaps. For example, a teacher created a protocol for her students to use when engaging in Socratic Seminars. The protocol addressed students' participation norms and sentence starters to use when discussing their text. This protocol was not given to the teachers by the Pathway Project facilitators. Other tools were created to address content standards, to expand concepts, and to scaffold students even more than what was covered by the Pathway Project curriculum.

Appropriation as Engaging Students

Finally, many teachers were concerned about the complexity level of the materials disengaging students so incorporated technology, created new tools to capture student interests, and developed games incorporating the Pathway Project tools for students. For example, another teacher developed a matching game for students to identify and define what the cognitive strategies were. Other teachers incorporated the cognitive strategies into silent reading activities or when discussing other types of writing.

Discussion

The students of the 17 teachers in this study incorporated many aspects of the Pathway Project curriculum and tools in their teaching. Many of them latched on to the tangible tools they were given such as cognitive strategies posters, book marks with sentence starters, and color-coding strategies that helped their students recognize gaps in their writing if they lacked textual evidence or commentary. What also resonated with the teachers was a process approach to writing and the meta-cognitive ways to engage with the texts they use in their curriculum. Thus, the teachers learned how to teach their students through appropriating the various tools from the Pathway Project. Separate analysis of the students' writing from these 17 teachers showed improvement in terms of the quality of their writing and quantity of their writing. In this regard, the teachers' appropriation of the Pathway Project tools led to an improvement in student learning around analytical writing. However, due to the limitations of the data we cannot attribute this improvement to just the Pathway Project curriculum. More systematic collection of data is needed to understand what aspects of the Pathway Project tool influenced the teachers' teaching and what other parts of their professional lives impacted their teaching. This limitation demonstrates the complexity that is learning from professional development.

Conclusion

The teachers appropriated the Pathway Project tools for different purposes and demonstrated the complexity of teacher learning through their enactment of the Pathway Project tools. The teachers became a Pathway Project teacher through their participation across two years. The teachers had to makes sense of the tools they were given and adopt and adapt them as they saw fit to meet their students' learning needs and to meet the needs of other professional development initiatives. The findings from this study demonstrate how complex teacher learning is and how professional development can trace that learning. Moreover, by using appropriation as one component of this complex endeavor, and demonstrating the varied ways teachers do so, this study moves the field beyond the levels of appropriation to consider the purposes of appropriation and the affordances and limitations of such appropriations on students' learning.

Teachers are the main vehicle in which education policy is enacted. Without attending to their learning, such as how they appropriate tools from professional development, sustained improvements in education cannot occur. Moreover, systematic understanding of the impact professional development has on different domains is also needed. Teaching is a complex practice. Research needs to acknowledge the complexity of teaching and recognize that just bringing teachers together is not enough. Rather, a holistic and comprehensive view of

teaching is needed. This study adds to our understanding of the impact of professional development on teachers, their learning, and ultimately their students' learning as a complex endeavor that is worthy to be studied through careful analysis of teachers' appropriation of tools and the factors that influence this process.

References

Baker, E.L., Barton, P.E., Darling-Hammond, L., Haertel, E., Ladd, H.F., Linn, R., Ravitch, D., Rothstein, R., Shavelson, R.J., & Shepard, L.A. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: EPI Briefing Paper.

Ball, D. L., & Forzani, F. M. (2011, Summer). Building a common core for learning to teach, and connecting professional learning to practice. *American Educator, 35*(2), 17-21.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Borman, G.D. (2005). National efforts to bring reform to scale in high-poverty schools: Outcomes and implications. *Review of Research in Education, 29*, 1-27.

Brown, A.L. & Campione, J.C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble& R. Glaser (Eds.), *Innovation in learning* (pp. 289-326). Hillsdale, NJ: Erlbaum.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd edition). Los Angeles, CA: SAGE Publications.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives, 8*(1), 1-44.

Darling-Hammond, L., Wei, R.C., Andree, A., Richardson, N., & Orphanos, S. (2009). State of the profession: Study measures status of professional development. *Journal of Staff Development, 30*(2), 42-50.

Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Towards better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199.

Desimone, L.M., Porter, A.C., Garet, M.S., Yoon, K.S., & Birman, B.F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81-112.

Echevarria, J., Short, D., & Powers, K. (2006). School reform and standards-based education: A model for English-language learners. *The Journal of Educational Research, 99*(4), 195-211.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.

Goldenberg, C., & Gallimore, R. (1991). Changing teaching takes more than a one-shot workshop. *Educational Leadership, 49*(3), 69-72.

Grosman, P., Smagorinsky, P., & Valencia, S. (1999). *Appropriating conceptual and pedagogical tools for teaching English: A conceptual framework for studying professional development*. National Research Center on English Learning and Achievement, Report #12011.

Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership, 59*(6), 45-51.

Hawley, W. D., &Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 127-150). San Francisco: Jossey-Bass.

Hill, H.C., Blunk, M. Charalambous, C., Lewis, J., Phelps, G. C. Sleep, L. & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430-511.

Kazemi , E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts. *Journal of Teacher Education, 59*(5), 428-441.

McLaughlin, M.W., & Talbert, J.E. (2006). *Building school-based teacher learning communities: Professional strategies to improve student achievement*. New York, NY: Teachers College Press.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd edition). Thousand Oaks, CA: Sage Publications.

Morris, A. K., & Hiebert, J. (2011). Creating shared instructional products: An alternative approach to improving teaching. *Educational Researcher, 40*, 5-14.

Neuman, S.B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532-566.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). 'How large are teacher effects?' *Educational Evaluation and Policy Analysis, 26*, 237-257.

Olson, C.B. (2011). *The reading/writing connection: Strategies for teaching and learning in the secondary classroom* (3rd edition). Boston, MA: Pearson, Inc.

Olson, C.B., & Land, R. (2007). A cognitive strategies approach to reading and writing instruction for English Language Learners in secondary school. *Research in the Teaching of English, 41*(3), 269-303.

Opfer, V.D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research, 81*(3), 376-407.

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Pea, R. (1993).Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed Cognitions: Psychological and Educational Considerations*, (pp. 47-87). Cambridge: Cambridge University Press.

Penuel, W.R., Gallagher, L.P., & Moorthy, S. (2011). Preparing teachers to design sequences of instruction in Earth systems science: A comparison of three professional development programs. *American Educational Research Journal, 48*(4), 996-1025.

Remillard, J.T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

Richardson, V. (1996).The role of attitudes and beliefs in learning to teach. In J. Sikula, T. Buttery, & E. Guyton (Eds.), *Handbooks of Research on Teacher Education* (pp.102-119). New York: Simon & Schuster Macmillan.

Richardson, V., & Placier, P. (2001). Teacher change. In V. Richardson (Ed.) *Handbook of Research on Teaching* (4th Edition, pp. 905-947). Washington, DC: American Educational Research Association.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press.

Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J. V. Wertsch, del Rio, P., & Alvarez, A. (Eds.), *Sociocultural studies of mind* (pp. 139-164). New York: Cambridge University Press.

Schoenfeld, A.H. (2011). Toward professional development for teachers grounded in a theory of decision making. *ZDM Mathematics Education, 43*, 457-469.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-21.

Stipek, D.J., Givven, K.B., Salmon, J.M., & MacGyvers, V.L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 213-226.

Talbert, J.E., & McLaughlin, M.W. (1994). Teacher professionalism in local school contexts. *American Journal of Education, 102*(2), 123-153.

Thompson, J., Windschitl, M., & Braaten, M. (2012). Developing a theory of ambitious early-career teacher practice. *American Educational Research Journal*. doi:10.3102/0002831213476334

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teacher and Teacher Education, 24*, 80-91.

vonZastrow, C. (2010). Making professional development investments that matter. *Learning First*. Retrieved from http://www.learningfirst.org/making-professional-development-investments-matter

Wallace, M. (2009). Making sense of the links: Professional development, teacher practices, and student achievement. *Teachers College Record, 111*(2), 573-596.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research, 79*(2), 702-739.

Wenger, E. (2010) Communities of practice and social learning systems: the career of a concept. In C. Blackmore (Ed.) *Social Learning Systems and communities of practice*. Springer Verlag and the Open University.

Yoon, K.S., Duncan, T., Lee, S.W.-Y., Scarloss, B., & Shapley, K. (2007).*Reviewing the evidence on how teacher professional development affects student achievement* (Issues &Answers Report, REL 2007 No. 033). Washington, DC: US Department of Education, Institute of Education Sciences

Learning in Low-Performing School Districts: Conceptual and Methodological Challenges Resulting from Network Churn

Kara S. Finnigan, Warner School of Education, University of Rochester, kfinnigan@warner.rochester.edu
Alan J. Daly, Education Studies, University of California-San Diego, ajdaly@ucsd.edu

Abstract: Turning around low-performing schools and districts has become a primary focus of educational policy across the country as a result of state and federal accountability policies implemented over the last decade. Our study uses theories of organizational learning and social network analysis to examine the structure and types of 'ties,' or relationships, among educators across an entire low-performing district. Our paper uncovers critical aspects of the context, some of which are directly related to the accountability policies that are meant to bring about improvement, which limit 'learning' in these schools and districtwide. Our study has important implications for both the understanding of learning processes in districtwide improvement, as well as methodological strategies for examining these.

Focus and Significance

A national push for higher levels of performance and accountability through federal policies and programs has increased the pressure on the schools and districts in the most challenging circumstances. Recent federal policies like NCLB and Race to the Top rely heavily on local school districts and their low-performing schools to engage in reform efforts to bring about improvement. However, most schools have struggled rather than improved. U.S. Secretary of Education Arne Duncan brought attention to this in his testimony to Congress when he noted that NCLB, "has created a thousand ways for schools to fail and very few ways to help them succeed" (Duncan, 2011).

Improving underperforming schools is challenging work that requires close attention to internal conditions in these schools, including the social relationships that facilitate improvement. Underperforming schools tend to be turbulent organizations with high staff turnover, multiple and changing reforms, and challenges related to leadership (Daly, 2009; Daly & Finnigan, 2011, 2012; Finnigan, 2010, 2012; Finnigan & Stewart, 2009), teacher quality (Sunderman, Kim & Orfield, 2005), and teacher motivation (Finnigan & Gross, 2007). Research in other fields has found that system-wide improvement is closely linked to the quality and structure of organizational relationships (McGrath & Krackhardt, 2003; Tenkasi & Chesmore, 2003) with frequent interactions supporting the transfer of tacit, non-routine, and complex knowledge, thereby allowing for collaborative problem solving and systemic change (Hansen, 1999; Reagans & McEvily, 2003; Uzzi, 1997). Within the field of education, research has found that schools with collaborative or trusting cultures are more likely to show signs of improvement and innovation (Bryk & Schneider, 2002; Mintrop, 2004; Mintrop & Trujillo, 2005; Moolenaar, 2010; O'Day, 2004). Beyond the importance of internal processes and relationships, much recent literature has found that greater attention must be paid to the larger district in which low-performing schools reside given the key role of the central office and the importance of a systemwide approach to improvement (Burch & Spillane, 2004; Datnow & Castellano, 2003; Honig 2006; Marsh et al, 2005; Smylie, Wenzel, & Fendt, 2003).

Building upon this prior work, drawing on the theoretical lens of organizational learning, and utilizing the methodological approach of social network analysis, our exploratory study examines schools and their larger district context as they attempt to improve under accountability policy sanctions. We sought to answer the following questions: What are the structure of relationships for leaders in low-performing districts and do these change over time? What are the structure of relationships within low-performing schools and how do these facilitate or hinder learning? To what extent do leaders in low-performing districts have the cross-sector connectedness and reciprocal ties necessary for large-scale learning and improvement?

Our study makes a unique contribution to the research because it involves longitudinal network data of leaders in low-performing districts. In addition, it focuses specifically on the relationships among and between both school and central office leaders to understand the district as a larger organizational unit. Finally, the paper examines the structure and types of relationships necessary for organizational learning. In our prior work we focus on specific aspects of organizational learning processes; here, instead, we pull together theories of social networks and organizational learning to examine not only the existence and types of relationships of educational leaders, but also how these change over time, to understand organizational learning and improvement in low-performing districts.

 © ISLS

Theoretical Framework

Organizational Learning

Learning in education is often used to describe the process through which individual students gain knowledge or skills in school settings. However, 'learning' is also important at the school and district organizational levels, particularly in the context of reform. As schools and districts, "continue to face a steady stream of novel problems and ambitious demands" (Leithwood, Jantzi, & Steinbach, 1995, pp. 3-4), a deeper understanding of learning process in schools and districts that are under tremendous pressure to perform may be useful especially given the growing numbers of these systems. Our research builds upon the work of organizational learning theorists (see, for example, Argyris & Schön, 1996; Huber, 1991; Levitt & March, 1988; March, 1991) in better understanding these processes.

Organizational learning is the process of detecting and correcting problems to improve organizational effectiveness (Argyris & Schön, 1996). The process of accurately diagnosing the underlying issues facing an organization is one of the first and most crucial steps in an organizations ability to 'learn' and improve (Argyris & Schön, 1996; Collinson & Cook, 2007). This suggests that the process of organizational learning involves understanding the important elements of practice, as well developing the underlying beliefs that support practice. Learning in an organizational sense leads members to change both behaviors (Levitt & March, 1988) and norms (Collinson & Cook, 2007) through a deliberate, rather than haphazard, process (Fiol & Lyles, 1985). Recent research suggests that this learning orientation toward reform and change can have significant impact on improving organizational performance (Hubbard, Mehan & Stein, 2006; Knapp, 2008). However, this work is still very much in its infancy.

Learning involves refining theories through "single loop" learning or through "double loop" learning, which requires a more careful examination of underlying assumptions, values, and beliefs that result in the emergence of new Theories in Use. Single-loop learning is conceptualized as learning that remains within the current organizational paradigm. In other words, "how best to achieve existing goals and objectives and how to keep organizational performance *within the range specified by existing norms* [emphasis added]" (Argyris & Schön, 1978, p. 21). Double-loop learning involves examining, "incompatible organizational norms by setting new priorities and weightings of norms, or by restructuring the norms themselves together with associated strategies and assumptions" (Argyris & Schön, 1978, p. 24). This type of learning requires examination of underlying values or assumptions that at one time may have been supportive of organizational goals, but now inhibit the organization's ability to learn. A key distinction is that single loop learning refers to incremental or routine changes, while double loop learning refers to transformational or more radical change and innovation (Easterby-Smith, Crossan, & Nicolini, 2000).

While the level of learning has been a long-standing debate in the literature, many theorists believe that organizational learning is more than just the sum of individual learning and results in institutional memory at the organizational level (Easterby-Smith, Crossan, & Nicolini, 2000, p. 785). As Stoll (2009) points out, learning processes involve dialogue, allowing members of the community to connect, discuss, and debate. In essence, organizational learning is "embedded in the deeply held beliefs and shared conceptualizations that develop among members of the organization over time as particular understandings and practices evolve through unconscious and regular interactions" (Supovitz, 2009, p. 709). Organizational learning, thus, involves social activities or the social processing of knowledge (Bransford et al, 2009; Hubbard, Mehan, & Stein, 2006; Marks & Louis, 1999), as individuals within the organization develop and share new knowledge and tools that results in commonly held ideas or practices or collective learning.

A final important aspect of organizational learning relates to the way in which ideas or practices enter the organization or evolve. March (1991) argues that organizations require a balance between exploration (exploration of new knowledge or experimentation) and exploitation (refinement or utilization of existing knowledge). Furthermore, double loop learning would suggest that some degree of exploration (or search for new ideas and practices outside of the organization) occurs as the organization moves beyond current norms and practices. Building upon the work of Levitt and March (1988), Honig (2008) expands upon these ideas, arguing that the search or exploration process may involve scanning the external environment for ideas or bringing individuals with expertise into the organization. In fact, this flow of information into and throughout the organization is critical to organizational learning (Huber, 1991). Organizational actors 'incorporate' these ideas or approaches into practice, either formally or informally, and through a 'retrieval' process adopt these practices over time when faced with new situations.

Social Networks and Social Capital

As Lin (2001) points out, the common denominator across theories of social capital is the understanding that it consists of, "The resources embedded in social relations and social structure which can be mobilized when an actor wishes to increase the likelihood of success in purposive action" (p. 24). Social capital is concerned with the resources that exist in social relations (referred to as 'ties') between individuals as opposed to the resources

of a specific individual. It is the structure and quality of those ties that ultimately determines opportunities for social capital transactions and access to resources (Burt, 1992; Granovetter, 1973, 1982; Lin, 2001; Putnam, 1993, 1995). Two aspects of social capital, networks and trust, frequently appear in the social capital literature (e.g., Bourdieu 1986; Halpern 2005; Nahapiet & Ghoshal, 1998).

The first element, networks, is primarily focused on how an actor is embedded in social relations, which forms a patterned structure of relationships (Nahapiet & Ghoshal 1998). In a social network, individuals are embedded within relationships, and these relationships are embedded in larger subgroups that eventually form a social network. The role of networks has been implicated as both supports and constraints in the process of organizational change, learning, and improvement (Balkundi & Kilduff, 2005; Bartol & Zhang, 2007; Leana & Van Buren 1999; Mehra et al., 2006; Penuel et al., 2009; Weinbaum Cole, Weiss, & Supovitz, 2008). This literature suggests that the structure of social networks can support organizational goals by facilitating the flow of information between individuals and overcoming problems of coordination (Lazega & Pattison, 2001; Tsai & Ghoshal, 1998). Many scholars have identified densely connected networks as a critical source of organizational advantage (e.g., Adler & Kwon, 2002; Leana & Van Buren, 1999; Nahapiet & Ghoshal, 1998; Walker et al., 1997), as those social interactions provide opportunities to build trust and as such significantly add to an organization's ability to innovate through supporting risk tolerant climates (Tsai & Ghoshal, 1998).

The second element, trust, has been identified as one of the most important affective norms characterizing a community (Nahapiet & Ghoshal, 1998). Trust is based on interpersonal interdependence (Rousseau, Sitkin, Burt, & Camerer, 1998) and involves an individual's or group's willingness to be vulnerable to another party based on the confidence that the latter party is benevolent, reliable, competent, honest and open (Cummings & Bromiley, 1996; Hoy & Tschannen-Moran, 2003). Trust, as a social capital resource, has been associated with cooperation (Hoy & Tschannen-Moran, 2003) and group cohesiveness (Zand, 1997). High levels of trust have also been associated with a variety of efforts that require collaboration, learning, complex information sharing and problem solving, shared decision-making, and coordinated action (Bryk & Schnieder, 2002; Cosner, 2009; Tschannen-Moran, 2004; Tschannen-Moran & Hoy, 2000; Lin, 2001). Bryk and Schneider (2002) suggest that trust is especially important for organizations that operate in turbulent environments, which is certainly the case in underperforming urban schools and districts. In essence, a predictability of relations gained through reciprocal interactions decreases the vulnerability between individuals as well as potentially increases the depth of exchange due to a willingness to engage in risk taking (Larson, 1992; Uzzi, 1997). In support of this claim, research suggests reciprocal as opposed to asymmetric relations provide mutual benefit to the relationship in effect creating a reinforcing effect (Lin, 2001). Reciprocated relations are, therefore, important in providing opportunities to build and deepen the norms of trust necessary for the exchange of reform related resources. Reciprocated relations provide opportunities for individuals to interact and learn together, which is important in educational systems oriented toward learning (Honig, 2008; Lave & Wenger, 1991; Wenger, 1998).

Through these two theoretical lenses, our study seeks to examine the ways in which underlying social networks of leaders facilitate or hinder organizational learning processes. It is the interaction between the activities and conditions of learning that that provide both a lens and process for organizational renewal, opportunities for learning, and better outcomes.

Methods and Data Sources

The study involves case study design (Yin, 2003) focusing on one urban district, which is a district 'in need of improvement' under NCLB, serving approximately 32,000 students. The district is 90 percent nonwhite, with 88 percent of students receiving free and reduced price lunches. Within the district, nearly all of the high schools and many elementary schools are identified as 'underperforming' based on state and federal accountability guidelines. We also include embedded cases of two high schools within this district. This district is an important case as it typifies many of the urban districts across the country that serve primarily students of color from low socio-economic communities, have a pattern of underperformance, and are engaged in district-wide improvement efforts to move off of sanctions.

The quantitative data collection occurred between 2010 and 2013 and involved a survey instrument administered to both school and district staff with both fixed-response items relating to organizational climate and technical aspects of organizational learning, as well as social network items. The organizational learning and climate items were developed based upon our theoretical framework, as well as incorporated items used in other contexts that were adapted to the school and district setting (see, for example, Garvin, Edmondson, and Gino, 2008) or used in schools (e.g., Tschannen-Moran and Hoy's (2000) trust scale. For example, our instrument involved questions relating to technical aspects of organizational learning, e.g., single/double loop learning and exploration/exploitation, as well as to the overall climate of the school. In addition, the survey instrument involved social network questions based upon prior network studies (Cross & Parker, 2004; Cross, Borgatti & Parker, 2002; Hite Williams, & Baugh, 2005) and targeted both instrumental (expertise) and expressive (vent) relationships. Respondents were asked to quantitatively assess a particular relationship with

each individual on a 4-point frequency scale ranging from 0 (not at all) to 4 (1-2 times a week). For example, regarding expertise ties respondents were asked the following: "*Please select the frequency of interaction for each school/district staff who you consider a reliable source of expertise related to your work*." The vent network was created based on the prompt, "*Please select the frequency of interaction with members of the school/district who you turn to when you need to vent*."

Each year, we administered the survey to the district's leadership team, which included 181 individuals over the 4-year period. We surveyed those in formal leadership positions in the district, including the Superintendent, Chiefs and Directors from the central office and principals at the school sites. During a three-year period we also collected data within two high schools that we use in this paper, surveying all educators in these schools including classroom and non-classroom staff. For both the school and district leadership team online surveys we used a bounded/saturated approach (Lin, 1999; Scott, 2000), meaning we listed all members of the particular group (school or leadership team) and respondents were not able to list any "outside" people that they connected to for example, teachers from other schools or clerical staff. The benefit of using this strategy is that it, coupled with high response rates, provides a more complete picture and more valid results compared with an unbounded approach (Lin, 1999; Scott, 2000). Response rates for the school and district level surveys range from 80 to 88%, thereby meeting the threshold for social network analysis (Scott, 2000).

We used SPSS to conduct the analysis of the survey items that related to organizational learning and climate and used UCINET software (Borgatti, Everett & Freeman, 2002), including Netdraw (Borgatti, 2002), for the social network analyses. Given that respondents tend to be more accurate at identifying ongoing patterns than determining occasional interactions (Carley & Krackhardt, 1999) and that we were interested in stable structural patterns (Krackhardt, 2001), we dichotomized the data for our analysis to include only the most frequent ties between actors, i.e., data indicating individuals interacted at least once every two weeks.

Based upon the different theoretical areas discussed above relating to tie structure and quality, we conducted a series of analyses to examine distinct network measures, such as *density*, the number of social ties between actors divided by the number of total possible connections, as a dense network is thought to be able to move resources more quickly than a network with sparse ties (Scott, 2000). We also examined *fragmentation*, the ratio of the number of disconnected pairs to the possible number of fully connected pairs within a network (Wasserman & Faust, 1994), and *centralization*, as a highly central structures allow a few members disproportionate influence over the flow of resources (Raider & Krackhardt, 2001). Finally, we examined *reciprocity*, or the proportion of mutual connections. We also examined differences relating to type of tie (instrumental versus expressive).

Major Findings

Our study uncovered three important findings that we discuss briefly. While we have four years of data that we can include in our presentation and final paper, due to space constraints for this proposal, we only include two years of network data in Figures 1-4 below.

Network Instability Undermines Learning

One of our main findings from the study relates to the conceptual and methodological challenges of network instability as school and district leaders leave their positions voluntarily or involuntarily from year to year. For example, from Time 1 to Time 2 only 55% of the same leaders were in these roles (we retained them in the sample even if they had moved positions within school and central office) across these two time periods. Given the critical importance of trust and the strong and collaborative relationships that result, this is extremely problematic for organizational learning, generally. In addition, the realities of the low-performing urban district are extremely challenging for using rigorous social network strategies. In essence, the ability to accurately examine changes in network structures and types of relationships over time is difficult when you have such a large degree of turnover from year to year – across a 4-year period it is nearly impossible. In the current study, we used matched comparisons across years to see the extent to which relationships changed in our other work. However, in some ways this misrepresents reality by only considering the relationships of those who stay and not representing all individuals in these networks. Here, instead, we show each year as a point time (not as a matched comparison) using one of our expressive or emotional network areas – who do you go to get the 'scoop' on the district. In the Figure, the black nodes are central office staff and gray are principals. As can seen across the two time periods, each year leaders in this district are having to re-establish underlying relationships, and some degree of connectedness among and between schools and central office at Time 1 is gone by Time 2, with a large proportion of leaders not connected to any other principal or central office staff as designated by the dots on the left hand side of the network map.

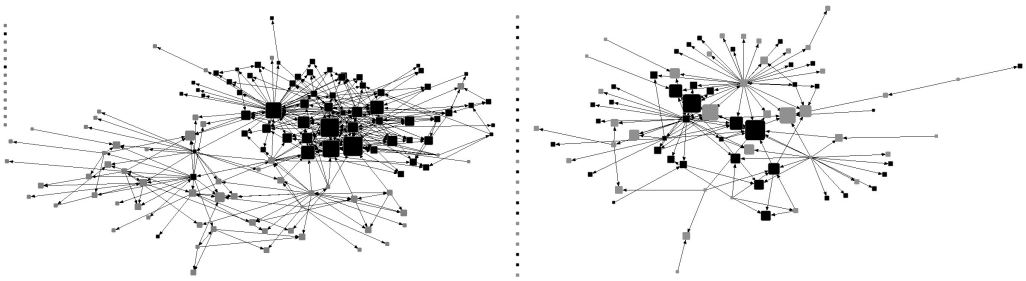


Figure 1. Instability of district-level leadership networks over time

Voluntary and Involuntary Movement of School Leaders Limits Exploitation

Our school-level network data indicates that administrators including both principals and assistant principals play and important role in the exploitation or search for new ideas through the sharing of research-based practices in schools. Given the low-performing nature of these schools having these administrators involved in exploitation, meaning bringing ideas from the external environment into the school, is critical to the learning processes. However, our data from these schools suggest that the larger challenging context along with the improvement strategies embedded in NCLB and more recent reform policies result in high levels of movement of these administrators. To illustrate this point we provide network maps from one of our schools at Time 1 and Time 2 as seen in Figure 2. In this case, the gray nodes are classroom teachers and black nodes are non-classroom teachers meaning administrators in the school, instructional coaches, counselors, etc. The nodes are sized by centrality so you can see the bigger nodes are the ones that most people go to for research-based ideas. The large central node in Time 1 is the principal of the school, who by Time 2 was moved out of the school involuntarily and into central office as part of the policy response to replace the principals of low-performing schools. The result of this move severely disrupted the sharing of research-based practices schoolwide as can be seen in the Time 2 map, with fewer ties and educators, particularly classroom-teachers relying on a different educator in the school with much lower centrality and far fewer ties. We found a similar result in our other high school although the assistant principal who was the source of research-based ideas in this school left voluntarily to move to principal position in another district.

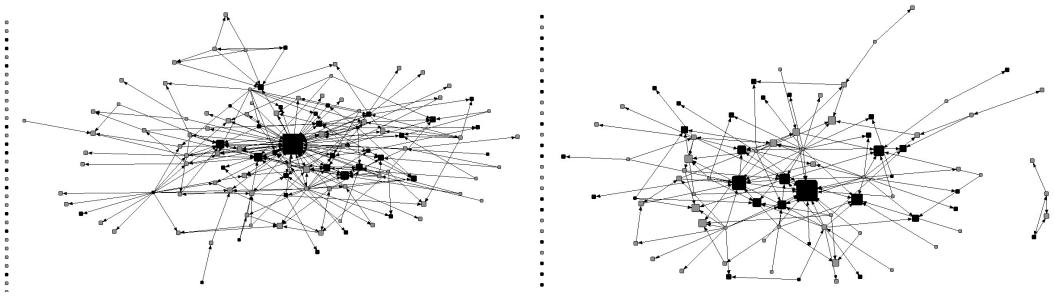


Figure 2. School level ties around research at Time 1 and Time 2

Lack of Reciprocal Relationships Necessary to Develop Trust and Exchange Ideas

A final area of importance is related to collaborative relationships, or reciprocal ties. Across both our school and district levels we found very few reciprocal ties when examining, in this case, the matched network maps over time. For example, although the number of ties for the same group of individuals increased from Time 1 to Time 2 for the instrumental or work-related tie relating to whom leaders turn to for expertise relating to their practice as seen in Figure 3, the proportion of reciprocal ties of all the ties that exist remained very low (under 20%). Furthermore, the expressive ties or more trust-related ties during the same time period decreasing suggesting that the conditions for reciprocal exchange were not sufficient in this district. Figure 4 illustrates the decreasing expressive ties from Time 1 to Time 2 and, as is evident in these maps given the scarcity of ties, few of these expressive ties are reciprocated.

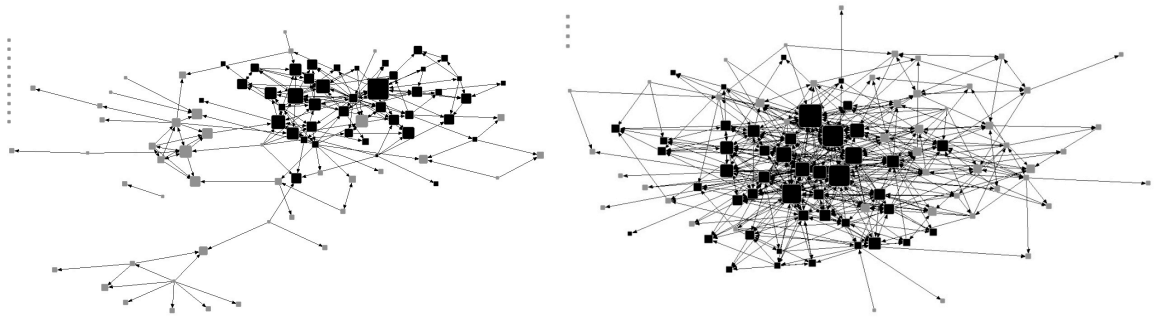


Figure 3. Increased but non-reciprocal instrumental ties over time

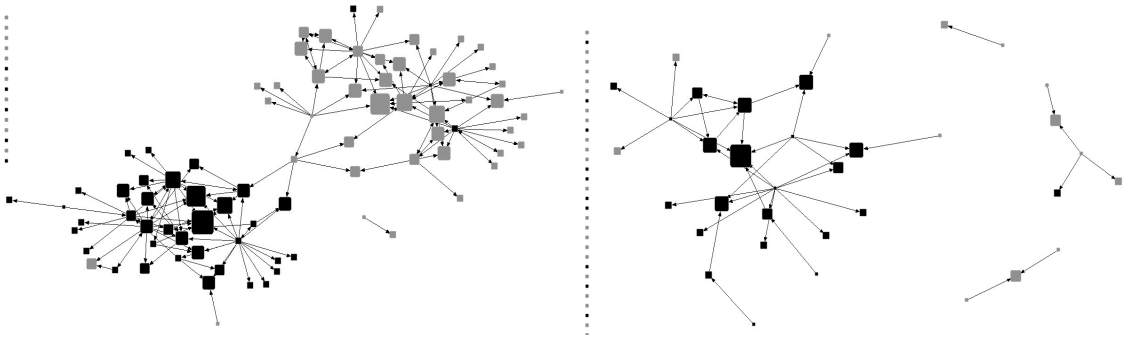


Figure 4. Decreased and non-reciprocal expressive ties over time

Conclusions and Implications

Using social network analyses techniques, we find that the network churn of school and district leaders creates an instability of relationships that undermines the potential for organizational learning. Perhaps connected to this network churn, we find few reciprocal relationships which are the cornerstone of communities of practice. Our study has implications for both the accountability policies that are driving reforms and improvement yet increasing the network churn at both the school and district level. First, the policies are implicated as they have increased levels of stress in these systems as the stakes become so high, resulting in high levels of movement in and out of the leadership team (including principal and central office). Second, they have directly caused some of this network churn through the school turnaround strategy requiring replacement of the principal based upon the number of years on sanction. While our findings have important implications for policy they also have implications for practice and research. At the district level, these data indicate that strengthening the trust within the system may need to be placed on the forefront of activities which can be difficult given the heavy emphasis on technical aspects of reform (e.g., around curriculum or testing). Finally, our study has implications for research given the changing and dynamic nature of these networks over time and strategies that are necessary to capture the underlying relationships, including structure of ties, given this network churn.

Relevance to Conference Theme

This paper contributes directly to the 11th Annual ICLS Conference theme, *Learning and Becoming in Practice*, in two ways. First, it focuses on how learning processes are situated within the practices of educators in low-performing school districts through an empirical study of the connectedness of these educators across the larger organizational context as well as within low-performing high schools. Second, it focuses on the practices for analyzing and modeling learning over time, paying specific attention to the conceptual and methodological issues that arise in these turbulent contexts that have not received sufficient attention in the theoretical or empirical research on organizational learning, school improvement, and accountability policies.

Endnotes

Both authors contributed equally to this paper.

References

Adler, P. S., & Kwon, S. (2002). Social capital: Prospects for a new concept. *Academy of Management Review,* 27(1), 17-40.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice*. Reading, MA: Addison-Wesley Publishing Company.

Balkundi, P., & Kilduff, M. (2005). The ties that lead: A social network approach to leadership. *The Leadership Quarterly*, 16, 941-961.

Bartol, K. M., & Zhang, X. M. (2007), Networks and leadership development: Building linkages for capacity acquisition and capital accrual. *Human Resource Management Review,* 17, 388-401.

Bourdieu, P. (1986). *The forms of social capital*. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241– 258). New York: Greenwood.

Bransford, J., Mosberg, S., Copland, M. A., Honig, M. A., Nelson, H. G., Gawel, D. Phillips, R. S., & Vye, N. (2009). Adaptive people and adaptive systems: Issues of learning and design. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 825-856).. Dordrecht, Netherlands: Springer International.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.

Burt, R. S. (2000). The network structure of social capital. In R.I. Sutton & B.M. Staw (Eds.), *Research in organizational behavior* (pp. 1-83). Greenwich, CT: JAI press.

Collinson, V., & Cook, T. F. (2007). *Organizational learning: Improving learning, teaching, and leading in school systems*. Thousand Oaks: Sage Publications.

Cosner, S. (2009). Building organizational capacity through trust. *Educational Administration Quarterly*, 45(2), 248-291.

Cummings, L., & Bromiley, P. (1996). Organizational trust inventory. In R. Kramer & T. R. Tyler (Eds.), *Trust in organizations* (pp. 302-330). Thousand Oakes CA: Sage.

Daly, A. J. (2009). Rigid Response in an Age of Accountability. *Educational Administration Quarterly*, Volume, 45(2), 168-216.

Daly, A. J. & Finnigan, K. (2012). Exploring the Space Between: Social Networks, Trust, and Urban School District Leaders. *Journal of School Leadership*.

Daly, A. J., & Finnigan, K. (2011). The ebb and flow of social network ties between district leaders under high stakes accountability. *American Education Research Journal*. 48, 39-79.

Easterby-Smith, M., Crossan, M., & Nicolini, D. (2000). Organizational learning: Debates past, present, and future. *Journal of Management Studies, 37*(6), 783-796.

Finnigan, K. S. (2012). Principal leadership in low-performing schools: A closer look through the eyes of teachers. *Education and Urban Society*, *44*(2), 183-202.

Finnigan, K.S. (2010). Principal leadership and teacher motivation under high-stakes accountability policies, *Leadership and Policy in Schools, 9*(2), 161-189.

Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation?: Lessons from Chicago's low-performing schools. *American Educational Research Journal*, 44(3), 594–629.

Finnigan, K.S., & Stewart, T. (2009). Leading change under pressure: An examination of principal leadership in low-performing schools. *The Journal of School Leadership*, 19(5), 586-618.

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *The Academy of Management Review, 10*(4), 803-813.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78*, 1360-1380.

Granovetter, M. S. (1982). The strength of weak ties: A network theory revisited. In P. V. Marsden & N. Lin (Eds.), *Social structure and network analysis* (pp. 105-130). Beverly Hills: Sage.

Halpern, D. (2005). *Social capital*. Cambridge: Polity Press.

Honig, M. I. (2008). District central offices as learning organizations: How sociocultural and organizational learning theories elaborate district central office administrators' participation in teaching and learning improvement efforts. *American Journal of Education*. *114*(4), 627-664.

Honig, M. (2009). What works in defining 'what works' in educational improvement. In G. Sykes, B. Schneider & D. Plank (Eds.), *Handbook of educational policy research* (pp. 333-347). New York, NY: Routledge.

Hoy, W. K., & Tschannen-Moran, M. (2003). The conceptualization and measurement of faculty trust in schools: The omnibus T-Scale. In W.K. Hoy & C.G. Miskel, *Studies in leading and organizing schools* (pp. 181-208). Information Age Publishing: Greenwich: CT.

Hubbard, L., Mehan, H., & Stein, M. K. (2006). *Reform as learning: School reform, organizational culture, and community politics in San Diego*. New York: Routledge.

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization Science, 2*(1), 88-115.

Knapp, M. S. (2008). District reform as organizational and sociocultural learning: How multiple theories shed light on a central leadership problem. *American Journal of Education 114*, 521-539.

Larson, A. (1992). Network dyads in entrepreneurial settings: A study of the governance of exchange relations. *Administrative Science Quarterly*, 37, 76-104.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Lazega, E., & Pattison, P. (2001). Social capital and social mechanisms and collective assets: The example of status auctions among colleagues. In N. Lin, K. Cook, & R. Burt (Eds.), *Social capital: Theory and research (pp. 185-208)*. New York: Aldine-de Gruyter.

Leana, C. R., & Van Buren, H. J., III. (1999). Organizational social capital and employment practices. *Academy of Management Review,* 24(3), 538-555.

Leithwood, K., Jantzi, D., & Steinbach, R. (1998). Leadership and other conditions which foster organizational learning in schools. In K. Leithwood & K. S. Louis (Eds.), *Organizational learning in schools* (pp. 67-90). Lisse, The Netherlands: Swets & Zeitlinger.

Leithwood, K., Leonard, L., & Sharratt, L. (1998). Conditions fostering organizational learning in schools. *Educational Administration Quarterly*, *34*(2), 243-276.

Levitt, B., & March, J. G. (1988). Organizational learning. *American Review of Sociology, 14*, 319-40.

Lin, N. (2001). *Social capital: A theory of social structure and action*. New York: Cambridge University Press.

March, J. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*(1), 71-87.

Marks, H.M., & Louis, K.S. (1999). Teacher empowerment and the capacity for organizational learning. *Educational Administration Quarterly*, *35,* 707-750.

Mehra, A., Dixon, A., Brass, D.J., & Robertson, B. (2006). The social networks of leaders: Implications for group performance and leader reputation. *Organization Science*, 17, 64-79.

Mintrop, H. & Trujillo, T. (2005). Corrective action in low-performing schools: Lessons for NCLB implementation from first-generation accountability systems. Education Policy Analysis Archives, *13*(48), 1-27. Retrieved June 29, 2007 from http://epaa.asu.edu/epaa/v13n48/.

Mohrman, S., Tenkasi, R. & Mohrman, A. (2003). The role of networks in fundamental organizational change. *Journal of Applied Behavioral Science*, *39*(3), 301-323.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantages. *Academy of Management Review*, 23(2), 242-266.

O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review, 72*(3), 1-31.

Penuel, W. R., Riel, M. R., Krause, A., & Frank, K. A. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record. 111*(1), 124-163.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6, 65-78.

Putnam, R. D. (1993). *Making democracy work*. Princeton, NJ: Princeton University Press.

Rousseau, D., Sitkin, S., Burt, R., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*(3), 393-404.

Scribner, J.P., Cockrell, K.S., Cockrell, D.H., & Valentine, J.W. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly, 35*(1), 130-160.

Stoll, L. (2009). Connecting learning communities: Capacity building for systemic change. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 469-484). Dordrecht, Netherlands: Springer International.

Supovitz, J. (2009). Knowledge-based organizational learning for instructional improvement. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 707-723). Dordrecht, Netherlands: Springer International.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal*, 41(4), 464-476.

Tschannen-Moran, M. (2004). *Trust matters: Leadership for successful schools*. San Francisco: Jossey-Bass.

Tschannen-Moran, M., & Hoy, W. K. (2000). A multidisciplinary analysis of the nature, meaning, and measurement of trust. *Review of Educational Research*, *71*, 547-593.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly, 42*(1), 35-67.

Weinbaum, E. H., Cole, R. P., Weiss, M. J., & Supovitz, J. A. (2008). Going with the flow: Communication and reform in high schools. In J.A. Supovitz, & E.H. Weinbaum (Eds.), *The implementation gap: Understanding reform in high schools* (pp. 68-102). New York: Teachers College Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge, MA: Cambridge University Press.

Yin, R. K. (2003). *Case study research, design and methods (3rd Ed.)*. Newbury Park: Sage Publications.

Zand, D. E. (1997). *The leadership triad: Knowledge, trust, and power*. New York, NY: Oxford University Press.

Shifts in Identification in a Hybrid Space

Kok-Sing Tang, National Institute of Education, Singapore, koksing.tang@nie.edu.sg

Abstract: In a hybrid space where people enact multiple identifications across time and space, this paper examines the question of why and how students shift from one identification to another in school. Through a design-based research in a high school physics classroom enacted to bring about a convergence of students' out-of-school discourses and school-based discourse, I analyzed the nature of identification undertaken by some students as they navigated multiple discourses. Using Bakhtin's work as an analytical frame, I suggest that shifts in identification should be seen as a temporary appropriation of a dialogic other's voice (or ideological stance) and suppression of one's preferred voice that is performed strategically according to one's situated interest at any particular point in time.

Introduction

In today's pluralistic and digital age, we are living in a hybrid world where we interact with multiple cultural systems and worldviews. We enact certain identifications (actions or behaviors that signal a recognizable affiliation; Lemke, 2009) within one discourse community and others within another community almost seamlessly in our daily life. In education, the tenet of exposing children to various disciplinary subjects is in essence expecting them to adopt multiple identifications affiliated to certain socially privileged communities. Yet, by the time most children enter formal schooling, they have already formed other identifications, some of which may present obstacles to their learning of a school-oriented identification. For example, to learn science is largely learning to be a scientist, and this involves learning its unique discourse (Gee, 1990), or way of interacting with or talking about the natural world (Lemke, 1990). However, there are also other ways of seeing and interacting with the natural world throughout human culture and history that constitute a different set of identifications (e.g., religion, sports). Given these multiple (and sometimes conflicting) identifications, how do students' shifts in identification take place?

This paper explores the above-mentioned question of identification shift in a hybrid space. Through a curricular intervention, a hybrid space in a classroom environment was fostered whereby the students' out-of-school discourses were directly juxtaposed with the official school science discourse. As explained by Barton and Tan (2009, p. 52), a hybrid space in a school setting is one where different discourses "coalesce to destabilize and expand the boundaries of official school discourse. Within this hybrid space, attention was given to the nature of identification undertaken by a group of students as they navigated multiple discourses, and examined why and how did some of them shifted one identification to another.

Theoretical Perspectives

Conventional thinking about identity tends to postulate a stable "essential self" (Erickson, 1968) with a durable set of psychological conditionings and social categories. This sets up an either-or distinction (e.g., expert vs. novice, scientist vs. non-scientist) that is problematic to our understanding of how people enact multiple identifications. Instead, the idea of plural shelves, which is increasingly more common, lends itself to our understanding of individuals enacting a range of identities in different contexts and for different purposes (Holland, Lachicotte, Skinner, & Cain, 1998; Lemke, 2009). While holding on to the latter model of identity, I am also interested in identity or identification shifts that occur in the moment and across timescales (Lemke, 2008) under the condition of hybridity. In this regard, I turn to Bhaha's hybridity and Bakhtin's voices to inform my work.

In his work on hybridity during colonialism, Bhabha's (1994) observes that whenever the colonizer sought to impose an essentialist discourse to shape the identity of the colonized (the Other) to become one of itself, it ended up producing something new to both the colonizer and colonized. Thus, he proposes the construction of a political hybridized subject that is "neither One nor the Other but something else besides, in-between" (p.219). Recontextualizing in educational setting, when a hegemonic disciplinary discourse like science is introduced in a classroom, Bhabha's insight provides a dynamic way of diffusing existing boundaries of established discourses and identities, and describes new possible spaces emerging from the interaction of multiple discourses.

While Bhabha's broad framework provides a new way of thinking about hybrid identities in theory, I turn to Bakhtin's *social voice* as a discursive lens for analyzing hybridity in practice. A voice is an ideological stance toward a discourse that is populated within an utterance. According to Bakhtin (1981), no utterance is completely unique and ideologically-neutral as people borrow and adapt others' voices in order to construct their own. The mixing of others' voices thus gives rise to *heteroglossia*, or the existence of speech diversity within a text or speech conversation. In other words, in any conversation, people inevitably mix the *dialogic*

other's voices into their own utterances as they respond to the preceding utterances and anticipate future responses from this dialogic other. As such, "all our utterances are filled with others' words [with] varying degree of otherness or varying degrees of 'our-own-ness'" (Bakhtin, 1986, p. 89).

Methodology

The data for this analysis are taken from a larger design-based research (Collins, Joseph, & Bielaczyc, 2004) in an Honors Physics classroom of a public high school, located in a predominantly white suburban community. In this research, a hybrid space was created through the enactment of a specially-designed curricular approach. Informed by the New London Group's (1996) pedagogy of multiliteracies, this curricular approach was enacted in the following procedure:

1. Before the start of a major curricular unit (e.g., mechanics), every student selected a text (henceforth *choice text*) he/she had read or was likely to read. While there was no restriction on the media (e.g., video, website) of this text, its content had to be related to both the student's interests and the physics unit they would learn shortly.
2. After collecting these texts, several strategies to harness the resources in the choice texts and address their differences with school-based texts were designed and implemented.
3. An assignment was then designed to guide the students in connecting their choice texts with the physics unit. The requirement was a four-page essay, which included two components. The first component was to explain a peculiar phenomenon behind the student's choice text "like a physicist" (i.e., using the language and concepts they had learned in class). The second component was to write a critical evaluation of how science was represented in their choice text in comparison to their textbook.
4. These procedures were then repeated for the next curricular unit (e.g., electricity).

Ethnographic methods (Spradley, 1980) were used to collect data in the classroom with the first author taking on the role of a participant-observer (e.g., co-teacher). The major data sources included daily videos and field-notes of classroom observations (70 lessons in total), 24 videos of student interviews, 3 teacher interviews, 63 students' choice texts, and 59 students' out-of-school texts/media. Qualitative data were analyzed using constant comparative analysis (CCA) and multimodal discourse Analysis (MDA). CCA is an inductive method that generates broad patterns and categories through a systematic comparison of specific incidents in the data (Glaser, 1965), and was used as preliminary method to organize and categorize the corpus of texts and interviews through the three stages of open, axial, and selective coding processes. This was complemented by the use of MDA to go in-depth into the discursive use of language and symbol systems in the textual, interview, and interactional data sources. The MDA approaches were informed by the traditions of SFL and social semiotics (Halliday, 1978; Lemke, 1995). Specifically, I used Lemke's (1990, 1995) semantic analysis to understand how different experiences and perception of phenomena are textually produced, Kress and van Leeuwen's (1996) "grammar of visual design" to examine how meanings are made with photographs and images, and critical discourse analysis (Fairclough, 1992) to relate intertextually these semiotic designs to larger societal and ideological underpinnings.

Findings and Analysis

From the analysis, I assert that shifts in identifications are necessary for students to navigate multiple discourses in a hybrid space. This identification shift involves an appropriation of a dialogic other's voice (or ideological stance) and a temporary suppression of one's preferred voice. However, the shift only occurs momentarily and the students perform this shift strategically in order to fulfil short-term goals.

For this paper, I present two cases to illustrate and support the above assertion. The first case involves a fifteen-year old girl whom I call Naomi. At the time of the research, not only was Naomi a high school physics learner, she was also a fan of science documentaries, a typical "A" student, an aspiring cardiothoracic surgeon, a cheerleader, and a devout Christian. For the curricular unit of electricity, Naomi selected an article from the website *answersingenesis.org*, which was introduced to her during a bible study class in her church. This article was written by an electrical engineer with a degree in Bible Theology to explain the electrical function of the human body (e.g., neural transmission) from a creationist perspective (see Savige, 1999). In a separate analysis, I have analyzed the multifaceted voices in this article, and how the voice of creationist design is dialogically opposed to that of random evolution.

In Naomi's evaluation of this article, she recognized the differences between the expectation of the article's targeted audience and that of scientists in general. Her dual identifications as a Christian and science learner were important in getting her to point out the different purposes and the "different angles" and ways of "presenting information" between the article and a science textbook. Interestingly, in her writing, despite being the targeted audience of the article, Naomi constantly referred to a generic "reader" and did not identify herself as a believer, as shown here:

> The author was trying to convince the *reader* that there was a creator behind the entire nervous system in humans, God... The author did a great job of supplying a simple explanation and opinion for an average *reader* who wants to know that there is a creator behind the electrical design in humans (*italics* added).

This deliberate shift in identification from the subjective position and immediacy of a limited first person perspective to a seemingly objective position of a distant third person perspective allowed Naomi to evaluate the differences between the creationist article and scientific texts, and was necessary in order for her to consider the point of view from scientific practices, which is an alternative voice in juxtaposition with the creationist voice she was comfortable with. Through this, Naomi wrote an explanation of neural transmission using the language and representations of physicists and biologists.

However, Naomi's appropriation of a scientific voice in her essay writing does not imply that she had permanently formed a new identification, and consequently resolved her personal conflict of evolution. In a separate interview with her, Naomi provided a telling example of how she managed the conflict and the different identifications expected of her. Despite Naomi's conviction that evolution was wrong, she managed to "pass evolution test with flying colors" and obtain an A overall for Biology. She explained that she needed an A because she "badly wanted" to be a doctor in the future. She also narrated that during the evolution test, she was frustrated and critiqued every question in her mind, while simultaneously giving the "correct" answer to each question. This suggests that in her identification shift to become an "A" science student and a prospective doctor, she suppressed external representations of her preferred voices (of creationism) consciously and temporarily for strategic reasons. Furthermore, it is unlikely that in this process, she would have to give up her religious beliefs and practices. In other words, in her dynamic negotiation of multiple discourses, she enacted her identification momentarily according to a calculated alignment between the situated circumstances and her projected goals.

In the second case, I focus on the identification shifts of two colorguard girls who were required to explain the phenomenon of colorguard tossing using the language of physicists. Colorguard is a popular extra-curricular activity in American schools and colleges where the participants use props such as poles, flags, and rifles to express dynamic movement in synchronization with the music from a marching band. An integral movement in a colorguard performance is a toss, which involves the spinning and throwing of a prop into the air with synchronized timing. This example includes a micro-genetic development of their talk with the researcher, and as such, illustrate more clearly the dynamic real-time appropriation of a dialogic other's voices in the students' utterances, and the identification shifts that ensued.

In a discussion between Evelyn and Lucy, the focus was on how physics principles were involved in a colorguard toss. The students started by explaining the process of tossing a colorguard flag and rifle into the air. In particular, Evelyn and Lucy wanted to explain how tossing a rifle and a flag were different. This explanation was mediated heavily through the situated use of their gestures (Roth, 2004), which is an important feature of colorguard Discourse. After listening to their explanations, the researcher asked them to explain in a different way by using a visual system:

Excerpt 1: Initial Tension

Turn	Speaker	Utterance & Non-verbal Actions
36	Researcher	Okay, you know what will help me a lot, and. and I think will help you as well is to. **draw what you just said**. And so I'm going to give you some pieces of paper.
37	Lucy	Right
38	Researcher	I want you to draw. whether is it a rifle or flag and. **draw some arrows** and. talk to each other, and I want to see how it goes. Okay?
39	Lucy	Right.
40	Evelyn	Okay.

As the researcher had difficulty in understanding the students' initial explanations, he asked them to "draw what [they] just said" (turn 36). This in essence was a bid for them to "textualize" (Bloome & Egan-Robertson, 1993) their embodied gestural actions into a free-body diagram, which is a specialized visual system of physics discourse. Two minutes later, the researcher left the group momentarily and Evelyn and Lucy made their sketches and commented on their drawings. At the same time, the video camera that was mounted in front of them continued recording.

With the discussion of colorguard removed from its typical event setting and situated in the physics classroom and recorded for the purpose of learning physics, Lucy in particular demonstrated an awareness of the need to employ terminology common to physics such as "height" and "rotation" instead of what is typically used in colorguard such as "gas" and "steering" to identify in with this new social ecology (Lemke, 2008).

Excerpt 2: Phenomenological identifications with new social ecology.

Turn	Speaker	Utterance & Non-verbal Actions
57	Evelyn	Okay. So.. this is my gas.
58	Lucy	Your gas?
59	Evelyn	Well, that's how I start (*inaudible*) right?
60	Lucy	Right. So essentially.. say which hand is the gas and which hand is the..
61	Evelyn	Okay, well. my push down is the **gas** (*pushed left hand downwards*), and this is my **steering** (*raised right hand*), so.
62	Lucy	Your **height**.
63	Evelyn	My.. (*laughed*) okay.
64	Lucy	Oh no, I'm like **I'm explaining it to the video**. (*pointed at camera*)
65	Evelyn	Oh.
66		(*Evelyn drew an arrow up on left side and arrow down on right side of her figure; see Figure 1, as Lucy watched*.)
67	Evelyn	Okay, this is.. my HEIGHT.. This is my. rotation. (*wrote height and rotation next to her sketch; see Figure 1*)

In this excerpt, a critical moment took place when Lucy demonstrated her keen awareness of the Bakhtinian dialogic other, even when this "other" was an inanimate camera directed at them. Due to the "presence" of this dialogic other, Lucy made a conscious effort to translate Evelyn's colorguard term (e.g., steering) into one that the teachers could understand (turn 62). In turns 66 and 67, Evelyn then self-corrected her own choice of words from a colorguard lexicon (e.g., gas, steering) to words more commonly used in physics (e.g., rotation, height). She then included those words in her sketch (see Figure 1).

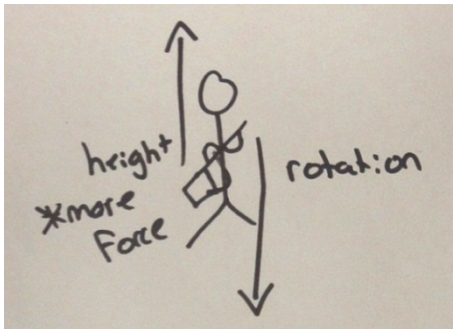


Figure 1. Evelyn's sketch of her bodily actions taken during a flag toss.

Immediately after this, Lucy, who had been watching her sketching, continued her momentary role as the teacher-facilitator in getting Evelyn to explain her actions for someone outside their colorguard discourse:

Excerpt 3: Momentary shift toward the expectation of the dialogic other.

Turn	Speaker	Utterance & Non-verbal Actions
69	Lucy	Right, so which hand uses more.. (*shook her head*) which hand. **has more force**.. if you are throwing a double?
70	Evelyn	(*Uttered a long sigh*) The height.

Turn	Speaker	Utterance & Non-verbal Actions
71	Lucy	Yah. (*smiled at Evelyn*)
72	Evelyn	Yah.
73	Lucy	Yah.
74	Evelyn	Yah. (laughing)
75	Lucy	Yah
76	Evelyn	(*Uttering to herself as she looked at her sketch*) This hand **has.. more force..** (*wrote more force on right side of diagram*) (*whispering*) **I don't know why..**
77	Evelyn	Cos if I. if it doesn't, then it's just gonna.. be too low
78	Lucy	Or.. Going to the ground if you use too much of the.. left hand than the other hand
79	Evelyn	Right.

Lucy's momentary role as the teacher-facilitator was evident from the I-R-E (Initiate-Response-Evaluation) she was enacting with Evelyn in the above excerpt. In turn 71, Lucy's followed up response of "yah" shows that the answer to her earlier question in this exchange structure (turns 69-71) was something that was already obvious to both of them. Evelyn's sigh in turn 70 and their back-and-forth strings of "yah" and smiles from 71 to 75 further show their mutual awareness that they were simply enacting this exchange for the dialogic other. This I-R-E exchange subsequently prompted Evelyn to say in (76) and write down "more force" next to the sketched right hand (see Figure 1). She then pondered why that was so and subsequently postulated a reason. After 8 seconds of looking at her drawing, she started to answer her own question by stating a possible outcome that it would "be too low" if less force was exerted (77). This was then followed by Lucy suggesting another probable outcome in turn 78. Throughout these two exchanges of I-R-E sequences (from 69 to 75) and self-question-and-answer (from 76 to 78), it was as though they were anticipating what the teachers would say to them if they were present. In other words, they dynamically shifted their identifications as they appropriated the voices and expectations of the dialogic other in their momentary utterances.

After the discussion and the completion of the essays, Evelyn and Lucy were interviewed to trace the development of their ideas to their essay writing. A key finding was how they were also able to momentarily suppress their preferred voices, styles, or stances in exchange for some incentives (e.g., academic grades). A case in point could be seen in how Lucy and Evelyn, in the course of fulfilling their assignments, had to adopt a reductionist stance to simplify their elaborate and colorful sport into abstract terminologies, symbols, dots, arrows, and numbers. Lucy for instance deliberately made her writing "boring" despite her self-identification as a creative writer as she desired a good grade for the assignment:

> It's boring because I just needed to like. I knew I have to get my point across without saying like. without getting into like describing. like it doesn't matter what color it is. It doesn't matter if it is fun or not. Like it's just. the information, and um... I was just being very basic like.. if I was to read this out loud, I will read this in a monotone… like even in the text I was using, she was like a little more fun with it, like 'she drops, the band director would commit murder' that sort of thing. But I was trying to avoid that, just because I know it was an educational text (italics added).

In Evelyn's case, she deliberately included more mathematical equations because she saw herself as "really good at math" and knew that one of the criteria for an A was to use relevant and accurate equations:

> When I think, I think very math-like, so this class is kinda hard for me because I know the math part, I do know pretty good the math, but the concept of it, it kinda trip me up. So like the scientific explanation, mine has a lot of equations in it, and it was mostly math, and I felt. I thought he [Brian] was going to be more like grade me a little bit higher because I knew I did so much math (italics added).

These examples highlight the active resourcefulness of the students in strategically shifting their identifications by momentarily appropriating a dialogic other's stance and temporarily suppressing one's own. Furthermore, they did so in order to fulfil their short-term goals as they navigated across the different discourses within a hybrid space.

Conclusion

This paper provides insights into how young people enact multiple identifications in a hybrid space. Contrary to a stable "essential self", my findings suggest that some youths could easily form hybrid identities that cater to the various discourses they find themselves in. I also found that such hybrid identities are not stable entities composed of a permanent concoction of two or more initial forms of identity. Instead, they are fluid states that are performed strategically according to one's situated interest at any particular point in time. This characterization is also supported by Gee's (2000) argument that today's well-off teens, whom he described as "shape-shifting portfolio youths", pick up a variety of experiences, skills, and achievements, (and I would add, voices and identifications), and are able to rearrange them dynamically and creatively for different circumstances.

One implication of this study is that instead of seeing identity or identification as a categorical frame that describes and confines people according to various social categories, it may be useful to consider identifications as resources or "toolkits" that people use to get by in their everyday life of negotiating multiple discourses. In this regard, our findings show that adolescents, by a certain age, have developed the ability to recognize and harness the appropriate identification toolkits to deal with the various discourses in schools. Nevertheless, an important qualification in this study is that these are relatively privileged suburban youths who have the social conditioning and dispositions (Bourdieu, 1984) to do well academically. What this would mean for other youths who lack the cultural capital and who may resist the identification shifts expected of them needs to be further explored.

This study raises several fundamental questions for future study and re-thinking of what is learning. First, how does a dynamic view of identification shift for strategic purposes inform or challenge current views of how people learn, particularly those that postulate learning as the exchange or transition from one stable form to another? To what extent is the calculated exchange of socio-economic goods (e.g., grades, career opportunities) lacking in constructivist theories of learning, and needs to be taken into consideration? Lastly, how can we change classroom research and practice as educators foreground students' agency in using identifications as dialogic resources to get by their everyday life, instead of seeing students as passive conformers of certain stereotypical social categories? These questions could help us reexamine our work of identifying how students identify themselves through an essential but daunting process that we call schooling.

References

Bakhtin, M. M. (1981). *The dialogic imagination: four essays*. Austin: University of Texas Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays* (1st ed.). Austin: University of Texas Press.

Barton, A. C., & Tan, E. (2009). Funds of knowledge and discourses and hybrid space. *Journal of Research in Science Teaching,* 46(1), 50-73.

Bhabha, H. K. (1994). *The location of culture*. London ; New York: Routledge.

Bourdieu, P. (1984). *Distinction : a social critique of the judgement of taste*. Cambridge, Mass.: Harvard University Press.

Fairclough, N. (1992). *Discourse and social change*. Cambridge, Mass.: Polity Press.

Gee, J. P. (1990). *Social linguistics and literacies : Ideology in discourses*. London [England] ; New York: Falmer Press.

Gee, J. P. (2000). Teenagers in new times: A new literacy studies perspective. *Journal of Adolescent & Adult Literacy,* 43(5), 412-420.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social problems,* 12(4), 436-445.

Halliday, M. A. K. (1978). *Language as social semiotic : The social interpretation of language and meaning*. London: Edward Arnold.

Holland, D., Lachicotte, W., Skinner, D., & Cain, C. (1998). *Identity and agency in cultural worlds*. London: Harvard University Press.

Lemke, J. L. (1990). *Talking science: language, learning and values*: Norwood, NJ: Ablex.

Lemke, J. L. (1995). *Textual politics : discourse and social dynamics*. London ; Bristol, PA: Taylor & Francis.

Lemke, J. L. (2008). Identity, development, and desire: critical questions. In C. Caldas-Coulthard & R. Ledema (Eds.), *Identity Trouble: Critical Discourse and Contestations of Identification* (pp. 17-42). London: Macmillan/Palgrave.

Lemke, J. L. (2009). Multimodality, identity , and time. In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 140-150). New York: Routledge.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review,* 66, 60-92.

Savige, C. (1999). Electrical Design in the Human Body. *Creation,* 22(1), 43-45.

Spradley, J. P. (1980). *Participant observation*. New York: Holt, Rinehart and Winston.

Shared Epistemic Agency and Agency of Individuals, Collaborative Groups, and Research Communities

Crina Damşa, University of Oslo, PO Box 1092 Blindern, 0317 Oslo, Norway, crina.damsa@iped.uio.no

Abstract: This paper aims to clarify the concept of shared epistemic agency, with its constituting aspects, and to examine research that illustrates how these are expressed in different settings that involve knowledge construction. Building on theoretical works from learning sciences, educational psychology and sociology, shared epistemic agency is conceived as a complex phenomenon that emerges in a dynamic way, and is defined as a capacity that enables groups to carry out collaborative knowledge-based activities that lead to a shared outcome. A comparative analysis of empirical studies of agency shows different ways epistemic agency is expressed in the context of individual and collaborative groups' learning and research communities' knowledge work. This discussion foregrounds the idea that creating intellectual interdependence, which is deemed necessary to co-construct knowledge, is an effort that can be assigned both to individuals and groups, but also to how the structural context affords and facilitates this interdependence.

Introduction

In recent years, constructive views of learning and knowledge work have put forward the argument that processes of knowledge construction require a particular type of engagement and conduct from participants (Bereiter, 2002; Valsiner & Van der Veer, 2000). In this context, the discussion about agency has emerged stronger; consequently, there is more interest in research concerning what agency means and how it is expressed. This paper addresses the notion of *shared epistemic agency* and how agency that is related to learning and knowledge work is being expressed at the levels of individual, collaborative groups, and research communities. I argue for the thesis that epistemic agency can have a shared nature and is expressed differently in various contexts. In support of this, I first conduct a conceptual discussion of shared epistemic agency, by analyzing the constituting aspects of the notion. Ultimately, I aim at a synthesizing definition of shared epistemic agency. Second, I provide examples of empirical studies, which investigated epistemic agency in the contexts of learning and knowledge work. This analysis is used to create a better understanding and to provide an illustration of how epistemic agency, and shared epistemic agency in particular, is expressed in different settings.

Before diving into the theoretical analysis, I present an example of two contrasting groups of students. Their characteristics were identified based on empirical analyses conducted in various research studies of collaborative learning.

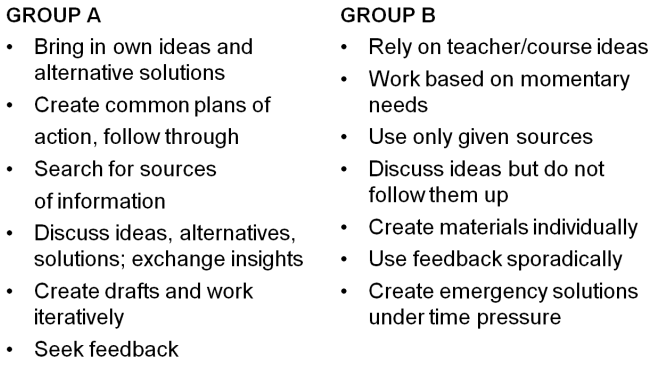


Figure 1. Contrasting collaborative groups

In all likelihood, this appears familiar to anyone who has been teaching in higher education. If we were to choose, probably the majority of us would prefer to work with groups such as Group A. While that is clear, have we thought in depth what is, in fact, happening in these groups? What makes the collaboration productive and meaningful? What drives it? Is it individual intellectual qualities, or the interest for the knowledge content, or for the final grade? I address some of these issues in the following discussion, with a particular focus on what drives these groups to engage in collaboration the way they do. I associate that with shared epistemic agency, which is viewed as the "capacity that enables groups to carry out collaborative knowledge-based activities that lead to a shared outcome" (Damşa, Kirschner, Andriessen, Erkens, & Sins, 2010, p. 154)

The "Question" of Agency

A quick look at the concept of agency, and how it has been theorized, shows that many of the present-day conceptions of human agency can be traced back to ideas of the Enlightenment. Original postulates referred to agency in terms of human freedom, in reaction to deterministic views of the religious morality of the time (Biesta & Tedder, 2006; Martin, 2011). Accordingly, people were conceived as willful beings, with capacity for independent judgment and autonomous action. *Philosophical* perspectives depicted agency in terms of an antagonistic relationship between a non-rational, normative action (based on Kantian ideas) and a rational–instrumental or utilitarian action (elaborated by American pragmatism and Continental phenomenology). Agency that emerges from the former reflects a moral will, in which actors (individual or collective) pursue free, moral action. On the opposite end, an individualistic conception pictures agency as rational but instrumental action. This is focused on agency serving an utilitarian purpose, the means to achieve an interest, or a material necessity. So, on one end, we have a focus on a self-legislating morality that can lead to good deeds, and on the other end, on purposeful thinking and action that serves the achievement of (individual) goals.

Within *sociology,* a vast discussion has unfolded with regard to whether agency is possible, how it should be conceptualized, and what are the conditions in which it emerges (see Figure 2 for a graphical representation of trends in sociology). Action, structure, and habitus were used as flagship notions. Discussions in the 70s and 80s commonly opposed agency to social structure and defined it as "the ability of actors to operate independently of the constraints of social structure" (Biesta & Tedder, 2006, p. 6). Later, Giddens (1991) and Bourdieu (1990) generated accounts that attempted to overcome this structure/agency dualism. Essentially, Giddens considered structure as a product of patterns of action, emerging from human activity but being constrained by rules. Bourdieu's view stressed that patterns, or habitual action, originate in the past, but then become structure and shape new, future activity and things in the world. Reactions to Gidden's view emphasized that agency is not just patterned action but action that breaks with patterns and defined rules and involves thought and reflexivity. A more recent conceptualization was put forward by Emirbayer and Mische (1998) in a seminal work within social theory. Essentially, these authors argue that the focus of previous conceptualizations is too much on judgment, routine, and structure, and not enough on the creative potential of agency.

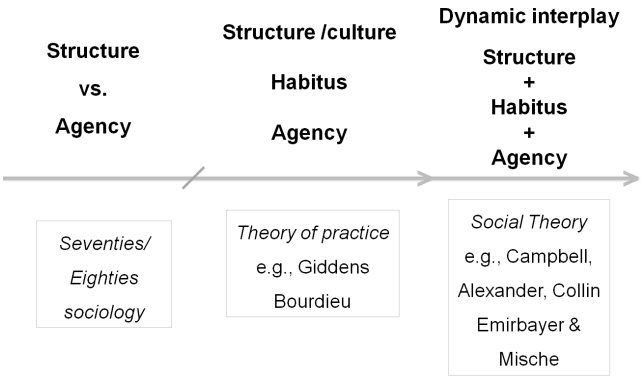


Figure 2. Sociological perspectives

A Relational–Pragmatic Perspective

To frame their conceptualization, Emirbayer and Mische state clearly that it is necessary to take distance from strict individualistic or holistic stances. The former are mainly represented by psychological theories, which focus on the ego as the main driver of decision-making and of the way to achieve an individual interest. The structural and cultural theories account for the latter. These emphasize the prevalence of structure and routinized action as following from this already established order. Emirbayer and Mische argue for a conception of agency that focuses on the dynamic interplay between routine, judgment, interest, and action. They further argue that agency can only be captured in its full complexity if it is situated within the flow of time and takes into account relational and structural aspects toward which participants can assume different orientations. Participants can be oriented toward the past, the future, and the present at any given moment, and while they move among these different contexts and timeframes, they also change their relationship to each other and to the context.

Following this re-conceptualization, an analytic distinction is possible between three dimensions of agency (Emirbayer & Mische, 1998). The *iterational dimension* is manifested in the participants' ability to recall, select, and apply taken-for-granted schemas of action developed through past activities. The agentic aspects do not simply lie in knowing such schemas but in how actors selectively recognize, locate, and implement such schemas. This refers mainly to how people capitalize on the existing body of knowledge and practices. The *projective dimension* indicates how agency implies orientation toward the future. This highlights the idea that people are able to challenge, reconsider, and reformulate their ideas, projects, and plans. This

assumes that people do not merely repeat past routines. They distance themselves from the established ways of action, can invent new possibilities for thought and action, and generate alternative responses to problems. It is about constructing images of "where they think they are going, where they want to go, and how they can get there (ibid., p. 984). Finally, the *practical–evaluative* dimension responds to the demands and contingencies of the present. This is about contextualizing experiences and knowledge to address problematic situations. It involves momentary judgment and deliberation and decision-making, about both means and ends of action. This capacity for practical evaluation enables participants to pursue their projects in ways that may challenge and transform the existing structures and action. Given this, agency is re-conceptualized "as a temporally embedded process of social engagement, informed by the past (in its habitual aspect), but also oriented toward the future (as a capacity to imagine alternative possibilities and outcomes) and toward the present (as a capacity to contextualize past habits and future projects with the contingencies of the moment)" (ibid., p. 963).

It is interesting to note that Emirbayer and Mische emphasize that an understanding of agency is only complete when it is able to account for the interplay between structure and agency. For them, agency is a dimension that is present in all empirical instances of human action, and there exist only actors who engage agentically with their structuring environments. This implies that all social action is shaped by the temporal–relational contexts of action *and* by the dynamic element of agency. This dynamic element itself leads to social action never being completely determined or structured.

Conception of Knowledge and Learning

In parallel, agency has emerged also as an interesting notion in the discussion concerning learning and knowledge work. The major epistemological shift that poses that knowledge is, in fact, not given but constructed assigned the individual and the environment completely different positions and roles in the knowledge processes. This shift and the discussion in itself brought about new concepts and the need to better understand the individual and the collective and their roles and interplay in the context of knowledge-bound activities. Within the sociocultural framework, Valsiner (1996) proposed a *bi-directional constructive* model that applies also to learning and knowledge. In this model, human actions and the social or cultural environment are given meaning by the participants through symbolizing activities. Accordingly, the individual is in an active process of relating to the environment and other individuals. The individual receives and transforms the information from/about the world into internalized personal knowledge, in the fashion Vygotsky (1978) described it. But the process is not unidirectional. Once the individual has constructed some form of personal knowledge, this becomes externalized in various forms—actions, artifacts, language; it then enters the communication with other individuals. Through this iterative process, knowledge is exchanged, adjusted, and elaborated. It is a co-construction process that can generate ideas, knowledge, and, ideally, development.

Commonly agreed upon, also by theoretical approaches such as the sociocognition or situated cognition, is that this co-construction process calls for conduct that renders possible the emergence of new ideas, insights, or knowledge (Greeno, 2006). Two aspects have been acknowledged as essential in this regard. One is the need for active involvement with the knowledge content, often framed in terms of productive engagement (Engle & Conant, 2002). Holding a belief about knowledge and simply memorizing given knowledge is a passive strategy, or lack of strategy. Participating actively involves re-constructing the meaning of this knowledge. An illustrative explanation is given by Bereiter (2002), in his elaboration of the belief and design modes in knowledge building, an approach that specifies deliberate activities for building knowledge in interaction. In a belief mode, learners attempt to understand given knowledge. In a design mode, a more participative and productive stance allows engaging with knowledge. Bereiter talks about *productive knowledge*, which allows the learners to use, question, and elaborate it, and proves as a stepping-stone toward new conceptualizations. A second aspect is what Valsiner and Van der Veer (2000) call "intellectual interdependence". Through this, individuals involved in shared contexts influence and guide each other. The individuals monitor one other's orientations and actions, modify their own intentions, and act in accordance. This generates the necessity of a relationship at the social-relational level. It is based both on social mechanisms of engagement, through means that mediate the interaction and on an individual's availability to be part of this relationship. Knowledge co-construction is realized through interaction and by using mediating means, among which communication with others is most important. In other words, this type of interactional achievement is realized in productive moment-to-moment interaction, in which a certain degree of intersubjectivity is required.

Intersubjectivity has been, traditionally, conceived as a collection of individual subjectivities or lifeworlds (Rommetveit, 1992), in the phenomenological sense, which describe a person's subjectively experienced world. He asserted that the social communication process starts from the assumption of shared understanding based on individual understanding, moves toward overcoming the mutual misunderstanding and results in joint, novel understanding. Matusov (1996) argues, however, that intersubjectivity cannot be viewed as a set of overlapping subjectivities or understandings, but rather as the coordination of contributions in joint activity. According to Crook (2013), intersubjectivity involves bringing states of knowing and doing in coordination with another person's; it involves *togetherness*, as he frames it. One of the important questions

Matusov brings to the fore is "how can the diversity of individual goals, ideas, actions unite people in activity?" While using different ways to express it, both Matusov and Crook suggest three stages in the emergence of intersubjectivity. Searching for common background and mutual "mindreading" is a first stage and involves coordinating with others about common goals and interests. An intermediate stage is about creating common ground for engagement, based on explicit communication among the participants. This involves a shared understanding, but which is by no means equal to intersubjectivity. This prepares for the final stage, of joint activity. The common ground and shared states drive joint action toward an outcome. In this way, intersubjectivity becomes leverage for productive interaction and the expression of agency in a shared manner.

Agency of the Epistemic Kind

Agency appears as a suitable concept that can depict what drives individual and collective learning and knowledge work, and within the field of the learning sciences, it has been conceived in different ways.

Productive agency, coined by Schwartz (1999), emphasizes the means and the way to arrive to a product. The underlying principles are grounded in Marxist ideas, namely, that any activity can be considered as having a product of some sort. In learning contexts, the main intended products are increased learner knowledge and understanding. Also, individual self-efficacy beliefs that drive one's behavior toward achieving a goal inspired this notion. Then, *epistemic agency*; etymologically, the term epistemic refers to knowledge, therefore, epistemic agency is considered the type of agency that entails acquiring knowledge. In education, the term was coined by Scardamalia (2000), as a synonym for collective cognitive responsibility, in the context of knowledge-building activities. It implicates students' willingness to see themselves as members of a community and to take responsibility for their own learning, but also for the advancement of the community's knowledge. In accordance, students demonstrate their epistemic agency through goal-setting, self-evaluation, and long-range planning. Finally, Pickering (1995) put forward the notions of *conceptual* and *disciplinary* agency, as an expression of different facets of accountability in intellectual practices. *Disciplinary* agency involves the use of accepted methods and procedures that are established in the practice of the domain. When individuals use an established method, agency is turned over to the discipline. In carrying out the method, a mathematician makes what Greeno calls "forced moves", and the individual's agency is limited to performing patterned actions, according to accepted practice. *Conceptual* agency involves the individual making choices and judgments about the appropriateness of methods and interpretations. Mathematicians can exercise conceptual agency when they engage in decision-making, exploration, and strategizing; they carry out *free moves*, in Greeno's terms. The conclusion that emerges from this quick overview is that most of these forms of epistemic agency have, in the strict sense, a focus on the individual. Eventually, Schwartz states that "agency develops through interaction, not only action" (2005, p. 50). But while the claim is that the collective is important in how the agency is played out, the individual expression of agency is prevalent.

These conceptualizations tend to take distance from the individualistic view and assign more importance to how knowledge is constructed and emerges through interaction with the structure, environment, and peers. But the notion of epistemic agency alone does not capture this shared nature in its full complexity. Conceiving shared epistemic agency as a construct based on intersubjectivity, which emerges from intellectual interdependence among individuals, seems paramount. Shared epistemic agency expresses this intersubjective nature, together with the productive features that lead toward knowledge being materialized into tangible, shared knowledge objects. Furthermore, it captures the dynamic interplay between individual agency and the surrounding structure, as it happens within a temporality that also influences how this agency unfolds and is being expressed. It becomes an evolving, constructed entity.

Agency in Different Settings

Individual Agency

In the second half of the previous century, the majority of the approaches in research on learning emphasized the individual cognitive aspects, the way learners could become efficient in individually processing information, and in monitoring and regulating this process (Bereiter, 2002). The concept of agency was framed within this theoretical context, with an emphasis on pragmatic aspects of the process. Ideally, the cognitive activity was to be organized through clear goal orientation and self-regulation. This is the closest we get to epistemic agency, which was equaled to self-efficacy and was assessed from a normative perspective. The unit of analysis in these studies was the individual belief and action.

Within this theoretical context, Bandura's (2001) theory of social cognitive learning generated studies on self-efficacy and self-regulation of learning at various educational levels. This research was strongly focused on the qualities necessary for a learner to be efficient and successful in control of his own knowledge acquisition process. In studies such as by Pajares (1996), agency was seen as being expressed through clear goal-orientation and self-efficacy beliefs in relation to the academic performance. Bandura (1990) identified agency through acting upon intentions and reflecting on action and Boekaerts and Corno (2005) through strategies of self-

regulation in classroom learning. Another line of research was generated by the self-determination theory, in which agency is equaled to autonomy. Deci and Ryan (2000) considered goal and action as insufficient for defining agency; values attached to beliefs and actions were needed too. Epistemic agency here is aimed at achieving individually set goals and places cognitive performance at the core of the notion, with much lesser attention given to knowledge and how that plays a role in the process.

Epistemic Agency: Individual within Collective

It is essential to acknowledge an area of research that situates epistemic agency between the individual and collaborative planes. This research accounts for the social but still allocates agency to the individual. The individual is expected to contribute to the community's knowledge or some collective outcome. This could be called the individual–collective perspective. The unit of analysis is the individual action but in natural connection to the collective activity.

A relevant set of studies follows Scardamalia's (2002) conceptualization of epistemic agency in knowledge building, wherein participants must contribute ideas to the advancement and improvement of the community's knowledge. Epistemic agency is defined as the responsibility taken by individuals for contributing to a collective set of ideas and to refining this knowledge through notes and comments. Most of the studies were conducted in the context of classroom-related inquiry-based learning. Epistemic agency was followed as the way for learners to direct and sustain their contribution to the pool of knowledge generated by the class. A study by van Aalst and Chan (2007) on the knowledge-building activities of secondary school students interpreted epistemic agency as being about further inquiry for understanding, since the ideas produced needed to be refined in order to be a meaningful contribution. Following knowledge-building ideas, Hakkarainen and colleagues (2004) designed and investigated networked learning. Their thesis was that agency is collectively constructed by individual actors that build and maintain an epistemic network. Students who share, facilitate sharing, and inquire further act as epistemic agents, who pursue collective epistemic goals and take responsibility for collective knowledge advancement. Analyzing students' social networks, Palonen and Hakkarainen (2000) concluded that students considered to be epistemic agents brokered knowledge and supported communication and information flow among members of the group. Also in this research, epistemic agency is placed in a collaborative space, but it is mainly expressed through individual actions that contribute to the collective goals.

Finally, studies on *disciplinary engagement* also touched upon the notion of agency of an epistemic nature. In Greeno's (2003) study of authoritative and accountable learning, the thesis is that participants express conceptual agency in domains, activity settings, and environments that facilitate that learning. This conjecture is consistent with analysis by Engle and Conant (2002), who identified aspects of productive disciplinary engagement in classroom discussions about biological concepts. Lastly, the notion of conceptual and disciplinary agency is transposed in a study of classroom competence in mathematics by Gresalfi and colleagues (2008). In their analysis of short episodes of discussion on mathematical concepts, agency is depicted based on whether a student initiates an idea, agrees with, elaborates on, questions, or disagrees with what someone else initiated, or refrains from responding. The most relevant conclusion is related to the role of the task, which can determine or influence students as they move productively through the task or resist being engaged in it. Again, agency is framed in relation to what that task involves in terms of content knowledge and strategies, at the individual level.

Shared Epistemic Agency of Collaborative Groups

A more specific conception of the interplay between individual and collective thought and action is provided by the sociocultural approach. As discussed previously, social input and individual action are intertwined, and this interplay can be empirically sought-after in the moment-to-to moment interaction. In general, this type of tight collaborative activity has been organized in small group learning or project-based work. In small group settings, the changes in orientation and action course, both of the individual and the group, are more detectable (Damşa & Ludvigsen, under review). As is assumed theoretically from this perspective, agency involves a close dynamic interplay between structural elements (of the instructional settings) and active individual participation; and also, between patterned activities and actions that might have creative value. The unit of analysis in this case is the group-level action, in which individual and collective actions are presumed to blend in naturally.

Charles and Shumar's (2009) study of student groups solving math problems in a virtual environment provides a good example of how shared agency can be enacted at the discursive level. The findings show an interplay between the individual and the group, in which group members managed their trajectory as a team. They also constructed meaning through dialog, bridged problem-solving episodes, and capitalized on jointly created artifacts. The authors think that the virtual chat, which is liberated from the social constraints of a physical space, encourages individuals to be agentic and to act like mathematicians. The epistemic aspect of agency is accounted for by the actions focused toward the mathematical domain—the shared aspect, through the intersubjectivity created by the group interaction in a coordinated manner.

Research by Damşa et al. (2010) and Muukkonen-Van der Veer (2010) focused on identifying and designing scenarios for collaborative construction of knowledge. Joint work aimed at generating new ideas and

elaborating and materializing them into shared knowledge objects was a central feature of the design. The type of agency considered to be necessary for such collaborative accomplishments is of a shared nature. Empirically, these studies searched for expressions of agency at three levels. First, in verbal interaction, agency was identified in the way groups raised awareness on a lack of knowledge, created shared understanding, and generated ideas. Second, they looked for how groups followed up on their verbal discussions and elaborations and used these in the construction of their shared knowledge objects. Third, these studies analyzed how groups made and followed up plans and how they coordinated their actions. The studies identified groups that displayed shared epistemic agency but also groups that encountered problems in both dealing with the knowledge content *and* their collaborative process. What follows is an example of verbal interaction that shows how a group discusses concepts that are central to their collaborative project. The excerpt below shows actions that indicate: lack of understanding, creating shared understanding, elaborating on an idea, and coordinating future actions. This is an example of how epistemic actions that indicate agency can be traced in momentary interaction of Group A.

Excerpt 1. Agency in momentary interaction (based on Damşa & Ludvigsen, under review)

1. Ann:	Can you explain? I don't really get it...	
2. Jane:	Erm, well, *in my school they have, for example, this common* ***assessment form*** *for the proof of competence*. All the assessors must use it, and the assessment of the same pupils by more **assessors** is discussed at the end among themselves. Isn't that something that has to do with **reliability?**	Sharing information
3. Liz:	While reading about validity, I came across things I recognized in practice. *The* ***task*** *for example, how that is covered by the tests, has to do with content* ***validity.***	Elaborating ideas
4. Ann:	*Also with* ***authenticity****. Tasks have to reflect the real work situation, the context...*	
5. Liz:	Yes, let's see how we do this! *We can each collect the information* from own schools, as an example. *We discuss it first again before we start writing*?	Coordination

When followed in time, throughout the course of the whole project, these actions can be indicative of a group's shared epistemic agency. There is a clear joint approach in this group's work, and data has the potential to showed how they consciously sustained this joint strategy when dealing with the knowledge involved and when organizing their process. This group's joint approach supported both productive collaboration that allowed knowledge co-construction and the emergence of shared epistemic agency.

Epistemic Agency in Research Communities

The theories underlining the conceptualization and the research of agency in knowledge communities are mainly of a sociocultural or sociomaterial nature. The individual is viewed as part of a structure, in which content, strategies, and methods are often customary. This system can involve collaborative practice, but it can also impede them. Agency can be expressed in the individuals' actions aimed at contributing to a) the knowledge domain or b) in them finding a way to navigate, in a productive manner, within the complicated structures of their domain practice. In general, the unit of analysis is individual action embedded in the larger structures or culture, but it can also be the collective action of a group within the community.

Studies by Knorr Cetina (1999) elaborate on epistemic agency as being part of the epistemic cultures emerging in different professional domains. In her empirical work, she analyzed how scientific groups function within particular epistemic cultures and touched upon agency as an aspect that plays a role in knowledge being constructed. Two of her laboratory studies, on the CERN atomic physics lab and the molecular biology lab, provide interesting illustrations of how scientific practice can develop. The analysis was focused on how physicists and biologists understand scientific work, the use of instruments and social structures (including organization of the lab). Three main conclusions can be drawn based on the findings. The *first* one is that epistemic agency contributes to the development and expansion of the domain's body of knowledge. The studies showed mainly how knowledge and epistemic objects emerge from scientific practices that draw upon individual and collective expertise. The knowledge generated in this way adds to the existing knowledge structures, in an incremental manner. This is the productive aspect of the epistemic agency, expressed on a scale of extended temporality. *Second*, Knorr Cetina shows that epistemic agency is formed in distinct ways in different epistemic cultures. She uses the notion of "machineries of knowledge construction" to depict the way actions and agency are shaped in a particular knowledge domain. It is a situated and contextualized process. In the case of the atomic physics lab, work was distributed in time and across sites. The characteristics of the methods within the domain of quantum physics also determined the work strategy. Epistemic agency was expressed in the way the

community kept the focus on shared goals and pursued that, despite the physical distance and other emerging problematic aspects. In the biology lab, where the group was smaller and the work took place in a confined space, agency was expressed more in relation to the social relationships and the material resources of the lab. And *finally*, Pickering's distinction between conceptual and disciplinary agency is applicable here. We see scientists working within the established boundaries of disciplinary practice but also some searching for alternative methods and strategies, attempting the free moves. But there is a certain specific logic to the interplay between agency and structure in this case. The structure here is represented by the existing knowledge domain and also by the set of scientific methods, procedures, and instruments. A scientist can choose to be compliant with the established methods of the domain and community or take a different path and bring in new ideas. The final conclusion is that "knowledge-centered practices" are forms of creative and constructive practices that go beyond routine and habits, and epistemic agency typically emerges when confronting non-routine problems.

Conclusions

This paper attempted to clarify the notion of shared epistemic agency, with its constituting aspects, and to examine empirical research that illustrates how these are expressed in different knowledge activities and settings. The theoretical analysis led to a depiction of shared epistemic agency as a complex phenomenon that emerges in a dynamic way. It can be viewed as a capacity that enables individuals, groups, or collectives to make appropriate judgments, to make plans and to pursue these through purposeful action, in order to achieve the construction of knowledge. A set of features emerged as characteristic of this notion. Productivity is expressed by following established ways of working with knowledge or by attempting new and creative strategies, methods, or interpretations. The sharedness refers to the social–relational aspects of the processes of knowledge co-construction. Theoretically, it is assumed that creating a strong intersubjective layer supports the manifestation of shared epistemic agency and the co-constructive process. Finally, the temporality refers to agency as an emerging entity, which unfolds through successions of intertwined thought and action. It combines the focus on past experiences and practices, with how these can be employed to attend to problems in the present and to create plans of action that aim at constructive processes projected in the future.

In practice, the studies analyzed showed that the shared epistemic agency is a complex construct, which is expressed empirically in different ways, depending on the context, its temporality, and the nature of the constructive activities. At the *individual level,* epistemic agency can be expressed in relation to individual work, but also in connection to collective work to which the individual might be expected to contribute actively. The purposive and productive aspects of agency are emphasized here, with the main focus on individual understanding of knowledge that is fed back into the collective knowledge. At the level of *research communities,* the iterational aspect of agency has a strong presence, since the knowledge and established practices are created through cumulative efforts and serve the community as a pool of options. Research communities are good examples of a dynamic interplay between structure and agency, since science and research are typically marked by creative–constructive actions geared toward the future. In order to build on the existing pool of knowledge, scientists might find creative ways to navigate within the structural context. *In collaborative groups*, shared epistemic agency is expressed in stronger terms. The way intersubjectivity is created can influence how agency unfolds in time and how its productive aspects are expressed. For example, when intersubjectivity is achieved, the joint resources bring about the intellectual interdependence that can afford sophisticated solutions and outcomes. One interesting point to make here is that technology mediation, as shown in Charles and Shumar's (2009) study, might open up new alternatives for shared agency to manifest itself.

Within this context, the interplay between structure and agency has specific characteristics. In institutionalized learning settings, individual students or groups are expected to acquire and master knowledge set by curricular demands. But they are also provided with a rich set of structural resources through the curriculum, which supports the knowledge construction process. Creative and alternative solutions are welcomed, but are not necessarily the standard way of work. In the research community, however, the stakes are higher when scientists engage in creative or innovative actions that break away from customary practices; but the rewards can be just as high. Individuals and groups that express epistemic agency and engage constructively with the knowledge domain and practice are the ones bringing science forward. In an ideal depiction, structure feeds into agency and allows the individual the freedom to choose a particular type of relationship with the contextual structure. Attempting to create intellectual interdependence is an effort that can be assigned both to individuals but also to how structural contexts afford and facilitate this interdependence. Regardless, understanding the notion of shared epistemic agency and creating the appropriate strategies to activate it is a complex challenge, to which further analytic work must provide the necessary input.

References

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1–26.

Bandura, A. (1990). Perceived self-efficacy in the exercise of control over AIDS infection. *Evaluation & Program Planning, 13*, 9–17.

Bereiter, C. (2002). *Education and Mind in the Knowledge Age.* Hillsdale, NJ: Erlbaum.
Biesta, G., Tedder, M. (2006). How is agency possible? Towards an ecological understanding of agency-as-achievement.Retrieved Oct 10 2013, from www.tlrp.org/project%20sites/LearningLives/papers/.pdf
Boekaerts, M., & Corno, L. (2005). Self-Regulation in the Classroom: A perspective on Assessment and Intervention. *Applied Psychology: An International Review, 54*(2) 199–231.
Bourdieu, P. (1990). *The Logic of Practice*. Stanford, Ca.: Stanford University Press.
Charles, E. S., & Shumar, W. (2009). Student and team agency in VMT. In G. Stahl (Ed.), *Studying Virtual Math Teams* (ch. 11, pp. 207–224). New York, NY: Springer.
Crook, C., (2013). Varieties of togetherness in learning – and their mediation. In M. Baker, J. Andriessen & S. Jaarvela (Eds.) *Affective Learning Together.* Routledge /Taylor & Francis Group: Oxford, UK.
Damşa, C. I., Kirschner, P. A., Andriessen, J. E. B., Erkens, G., & Sins, P. H. M. (2010). Shared epistemic agency – An empirical study of an emergent construct. *Journal of the Learning Sciences*, 19(2), 143–186.
Damsa, C.I. & Ludvigsen, S. R. (under review). *The collaborative construction of what? Learning through interaction and co-construction of knowledge objects in teacher education.*
Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.
Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology, 103, 962–1023.*
Engle, R., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement. *Cognition and Instruction, 20, 399–483.*
Greeno, J. G. (2003). Situative research relevant to standards in mathematics education. In J. Kilpatrick, W. G. Martin, & D. E. Schifter, (Eds.). A research companion to Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.
Greeno, J. G. (2006). Authoritative, accountable positioning and connected, general knowing: Progressive themes in understanding transfer. *Journal of the Learning Sciences, 15* (4), 539–550.
Giddens, A. (1991). *Modernity and Self-identity: Self and Society in the Late Modern Age. Stanford, CA:* Stanford University Press.
Gresalfi, M.S., Martin, T., Hand, V., & Greeno, J.G. (2008). Constructing Competence: An Analysis of Student Participation in the Mathematics Classrooms. *Educational Studies in Mathem*atics, *70* (1), 49–70.
Hakkarainen, K., Lonka, K. & Paavola, S. (2004) *How Can Human Intelligence Be Artificially Augmented Through Artifacts, Communities, Networks?* Paper at the Conference on Motivation, Learning, and Knowledge , IKIT and Karolinska Instituted, June, Stockholm, Sweden.
Knorr Cetina, Karin (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.
Martin, J. (2004). Self-Regulated Learning, Social Cognitive Theory, and Agency, *Educational Psychologist, 39*(2), 135–145.
Matusov, E. (1996). Intersubjectivity without agreement. *Mind, Culture, and Activity, 3(1), 25–45.*
Muukkonen-van der Meer, H. (2010). *Perspective on Knowledge Creating in Higher Education* (Unpublished doctoral dissertation). University of Helsinki, Finland.
Pajares, F. (1996). Self-efficacy beliefs in achievement settings. Review of Educational Research, 66, 543–578.
Palonen, T., & Hakkarainen, K. (2000). Patterns of interaction in computer-supported learning, In B. Fishman & S.O'Connor (Eds.), *4th Conference of the Learning Sciences* (pp. 334–339). Mahwah, NJ: Erlbaum.
Pickering, A. (1995).*The mangle of practice: Time, agency, and science.* Chicago, IL: Chicago University Press.
Rommetveit, R. (1992). Outlines of a dialogically based social-cognitive approach to human cognition in communication. In A.H. Wold (Ed.), *The Dialogical Alternative: Towards a Theory of Language and Mind* (pp. 19–44). Oslo: Scandinavian University Press.
Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a Knowledge Society* (pp. 67–98). Chicago: Open Court.
Schwartz, D. (1999). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.) *Collaborative learning: Cognitive and computational approaches* (pp. 197-218). NY: Elsevier .
Schwartz, D., & Okita, S. (2004). *The Productive Agency in Learning by Teaching.* Retrieved March 18, 2008, from http://aaalab.stanford.edu/papers/Productive_Agency_in_Learning_by_Teaching.pdf
van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences, 16, 175–220.*
Valsiner, J., & Van der Veer, R. (200*0). The Social Mind: Development of the Idea.* New York, NY: Cambridge University Press.
Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.
Wertsch, J. V. (1991). *Voices of the Mind: A Sociocultural Approach to Mediated Action*. Cambridge, MA: Harvard University Press.

Interrogating the Divide: A Case Study of Student Technology Use in a One-to-One Laptop School

Nicholas Wilson
University of Massachusetts, Amherst
Furcolo Hall
College of Education
Amherst, MA 01003
nwilson@educ.umass.edu

Abstract. This study examines the learning experiences of a cohort of students from historically marginalized backgrounds (ethnically- and socioeconomically non-dominant, as well as academically-underachieving) at a one-to-one laptop school to uncover institutional structures and teaching practices that contribute to the reproduction of digital education inequity. Using a sociocultural framework that incorporates activity systems theory (Engeström, 1987) this research reports on how tensions and contradictions between institutional, instructional, and student perspectives on learning in a one-to-one environment foreclose opportunities for agency and technology literacy development, in spite of access to 21st century learning technologies.

Purpose

The purpose of this study is to examine institutional structures and teaching practices that impact how students from non-dominant ethnic, socioeconomic, and academically underachieving backgrounds learn with technology in a one-to-one laptop environment. Specifically, the study investigates the types of activities students are asked to perform with technology, how such activities support the development of technology literacy skills and student empowerment, and finally, how students "come to terms with" (i.e., respond to, cope with, resist, or embrace) emergent tensions between their personal learning goals, and "what counts" as learning in the classroom.

An increasing number of studies on digital education inequity have linked a mixture of complex sociocultural influences to students' information communications technology (ICT) literacy development and the reproduction of the "digital divide" (Harris, 2010; Sims, 2013; Subramony, 2007; Warschauer et al., 2004; Warschauer & Matuchniak, 2010; Windschitl & Sahl, 2002), suggesting a growing need for research to qualitatively examine the nature of technology integration and teaching practices in our nation's schools – especially those in historically underserved communities. Whereas the most fundamental area of digital inequity remains access to computing technology and high speed Internet (Hohlfeld, Ritzhaupt, Barron, and Kemker, 2008), further divisions persist along the lines of frequency of technology use, the purposes for which students use technology, and students' capacity to utilize technology in personally empowering ways (Hohlfeld et al.). This study provides a microanalysis of students' practices with technology to examine the nature of contradictions that inhibit 1) the frequency of opportunities for students to learn with technology, 2) the use of technology for higher-order learning activities, and 3) how students are empowered to use technology for personal and academic growth.

Potential Significance of the Work

This study will contribute to existing literature on digital education inequity and technology integration. The central focus of technology's role in this study is imperative for understanding the nature of barriers to effective, technology-mediated learning. While socioeconomic factors hold an obvious implication for the disparity of technology skills observed across rich, poor, urban, and rural communities alike, recent studies suggest that other social factors play an important role in the development of technology literacy and 21st century skills. This study addresses the role of social practices and embedded institutional structures in the development of those skills, highlighting the tensions that students and teachers must navigate, bridge, or challenge to make technology integration successful and effective.

As Hohlfeld et al. (2008) and others have suggested, the wide disparity of technology use observed across our classrooms is a primary symptom of digital education inequity today. Indeed, the very existence of such inequity, especially in the face of reported increases in access, implies a disconnection between common conceptions about teaching, technology, and how integration translates into opportunities for learning. A study on tensions between structures and practices can contribute to our knowledge of these problems by examining the nature of social, technical, and relational processes that afford such learning opportunities. Instead of

simplified, ambiguous interventions like more professional development hours or the installation of more expensive equipment, research should seek to understand how technology can be used to construct knowledge, how that use is impeded or enhanced by social and structural practices.

Theoretical and Methodological Approaches

Conceptual Framework

Over the past two decades, researchers have observed persistent gaps in the development of technology literacy skills between students from underserved ethnic and socioeconomic backgrounds, and their more affluent, well-served counterparts. Criticizing oversimplified notions of the "digital divide" that attribute such gaps to a deficiency of access to personal computers and high-speed Internet, recent studies suggest, rather, that the problem is tied to longstanding social and educational inequities. Approaches to investigating student technology use in light of these suggestions has required a considerably more complex understanding of student technology use than those that have deterministically proposed that merely providing access to technologies will narrow the divide.

Borrowing from theories of social practice (Bourdieu, 1977) and "structuration" (Giddens, 1985), some scholars have begun to consider these "digital inequities" to be the result of myriad institutional and social influences that impact not only how students use technology for learning, but what skills, literacies, and cultural capital are even valued in the context of schooling (Ito et al., 2010; Ito et al., 2013; Mouza, 2008; Sims, 2013; Warschauer, 2004; Zhang, 2010). Indeed, these studies situate technology literacy as the product of the "systems of relations" (Lave & Wenger, 1991, p.53) between a dynamic constellation of individuals, members of the learning context, and the human- as well as material- resources that are available in the environment. This perspective effectively destabilizes the notion that technology literacy and learning are somehow fixed or natural outcomes of participation in technology-integrated, school-based exercises. Rather, they suggest, technology literacy development entails a continuously negotiated range of dispositions, between the individual student, and the sociohistorical context of the learning activity. As the constituent parts of these "activity systems" (Engeström, 1987), change from context to context, so do the identities and social positions of those participating in the activity. When considered through such a lens, technology literacy development, and the factors that enable or inhibit it, appears more closely tied to relationships of power, privileged social practices, and the "structuring structures" (Bourdieu) of institutions than mere access to 21st century learning tools. This study adopts a sociocultural approach to examine the technology-related classroom activities students partake in, and how their participation relates to institutional values of 21st century learning, in an effort to understand the mechanisms that reproduce digital education inequities.

Fundamental to this study is a conceptualization of digital education inequity as the result of the reproduction of social structures and practices that perpetuate and institutionalize the marginalization of non-dominant forms of knowledge, culture, and values (Ito et al., 2013). The valued elements of culture, some argue, are derived from traditions and practices that are based in certain ontological and epistemological assumptions, which have become ingrained into our ways of manipulating and seeing the world, through repetition and routinization over many generations (Bourdieu, 1977; Giddens, 1985). This "naturalization" of culture is the heart of the meaning of social practice and reproduction. Hence, social customs, such as teaching practices and "schooling" (Aronowitz & Giroux, 1993) are institutionalized ways of knowing and doing that reflect certain privileged values.

Arguably, the mechanisms of social reproduction at work today have evolved alongside the somewhat recent emergence of personal technologies, such as computers and cell phones. Increased access to these technologies has shifted our cultural topography to the extent that historically marginalized populations have increasing access to channels of cultural participation, as well as to means for constructing entirely new modes of participation. Yet within the education system, traditional notions of literacy and the "transcendent script" (Gutierrez, Rymes, and Larson, 1995) of schooling still dominate the current paradigm of classroom instruction. Indeed, some argue that the tools of reproduction used to separate the culturally-rich from the culturally-deficient (Bourdieu, 1977) (methods such as tracking, standardized testing, and other methods of quantifying "intelligence" and "ability" [McDermott, 1993; McDermott & Varenne, 1995; McDermott & Varenne, 1996]) in the classroom have become even more entrenched (Aronowitz & Giroux, 1993; Darling-Hammond, 2007). While cultural participation continues to evolve in its form and in its content, many of the technologies appropriated by communities and individuals to communicate, interact with, and produce knowledge are often suppressed or constrained in the classroom (Lemke, 2010). This begs the question of what educational practices and structures serve to alienate culturally active youth from their academic environment, and marginalize the quality and content of their cultural participation (Wang & Ching, 2003).

The current research follows a small cohort of high school students from non-dominant ethnic and socioeconomic backgrounds in a mixed-level introductory Biology class, and examines their experiences learning with technology as participants in the school's one-to-one laptop program. In this ethnographic

account, I investigate how the conditions of activity – institutional structures, the teacher's technology skills, attitudes, and beliefs about technology and instruction, the tools students use to accomplish various learning tasks, etc. – influence why and how students use technology in the classroom, and to what ends.

Methods

Data Collection

This study took place at a mid-sized high school in a working class town in eastern Massachusetts, dubbed "Bayside" (pseudonym). At the time of this study, the school was in the second full year of its one-to-one laptop program – an implementation that had already spanned seven years of research, planning, and development. Data collection for this study encompassed a number of strategies aimed at understanding the social and historical context of the research setting, the practices of individuals within that setting, and the tensions that emerged with regards to the role of technology use and students' dispositions towards technology and learning. Field notes of each class meeting were documented as I moved in between silent-observer and participant-observer roles throughout the year, occasionally helping students navigate unanticipated technical disruptions. Semi-structured interviews with the focal teacher and student participants took place throughout the year, centering on emergent themes related to the use of technology and teaching practices.

During the latter half of the year, in addition to documenting field notes, each class meeting was video recorded to provide a closer examination of the classroom environment, and a finer-grained analysis of the skills students employed for engaging in various technology activities. Finally, learning artifacts, including an entire archive of the course's online site, and official documentation related to the school's one-to-one laptop program rounded off this large corpus of data.

Data Analysis

Analysis of the data involved a constant comparative analysis (Glaser & Strauss, 1967; Strauss, 1987; Corbin & Strauss, 2008). Constant comparative methodologies stem from the foundations of grounded theory (Glaser & Strauss), which focuses on the identification of "categories, properties, and hypotheses that state relationships between categories and properties" (p.8), for the purpose of proposing theories that are interpretive in nature (as opposed to prescriptive). Data was scrutinized throughout the collection process for recurring themes and relationships, employing a phasic process, which included rounds of open coding, axial coding, and selective coding, concurrent with frequent cross-checks to establish validity between codes (Demetriadis & Kamberelis, 2006).

Anthony and Clark (2011) used a constant comparative analysis to situate the context of technology-related activities, identify barriers to students' technology use, and then categorize students' mechanisms for coping with those barriers. To do so, they borrowed from Engeström's (1987) theory of activity systems to investigate the nature of teachers' dilemmas of practice when integrating technology in a one-to-one laptop environment. In addition to interviews and surveys with key laptop program stakeholders, the researchers conducted an analysis of official documents that were a part of the laptop program, including the district technology plan, technology use policies, the school's technology vision statement, professional development plans, curriculum materials, student assignments, and other documents related to the program. The current study utilized similar methods to explore tensions and contradictions that affected students' use of technology.

First, field notes, interviews, video recordings, learning artifacts, and school documents were coded along a schematic of activity systems analysis to identify the objectives, mediating artifacts, rules, divisions of labor, and the "classroom microculture" (Barab, Barnett, Yamagata-Lynch, Squire, & Keating, 2002) that embodied the Introduction to Biology class's use of technology. From these codes, a more focused, micro-level analysis took place, identifying areas of tension or contradiction that appeared to affect 1) the frequency and purpose with which technology was integrated into instruction, 2) the role of technology within various technology-integrated activities, and 3) the creation of opportunities for students to learn about- and with technology. Finally, these tensions and contradictions were examined to identify the systemic issues that impacted the focal students' learning experiences in the context of a one-to-one laptop program.

Findings and Discussion

After the end of the school year, the focal teacher and I conducted our final interview to discuss the participant students' achievement in the class, the range of technology integration strategies he adopted over the year, and his perceptions of students' dispositions towards technology and learning. After the final exam, only three of the focal students received a passing grade for the year, one of which the teacher described as a "Gentleman's D-minus" – effectively a merciful act bestowed upon a senior whose final grade teetered on the edge of failure, and which put his graduation and enrollment in a local community college the following year at risk. One student from this group who did pass experienced a massive slump in grades during the second semester, "banking" good enough scores early on to still pass the course within minimal effort over the final months of the school

year. As we discussed these disappointing, but somehow all-too-common outcomes, we circled back to the purpose of the Bayside High School one-to-one laptop program, the role of technology in the Introduction to Biology course, and the tensions that impacted how, when, and why the students used technology in the class. While the teacher spoke of the difficulties integrating technology in a class of some thirty students, and the need for instructional resources to support the diverse range of learning backgrounds and differences in such an environment, what struck me the most was the account of his students' dispositions towards schooling, and the way these dispositions transcended their use of technology throughout the class. "Just get it done," was their mantra – a theme that pervaded these students' orientations towards technology and learning throughout the year.

As I combed through my field notes, interview transcripts, and video recordings for evidence to disprove the lamentably low bar that the students seemed to set for themselves, what began to emerge was the image of a systemic tension that is reminiscent of deterministic technology policies: the institutionalization of the very tools we as educators hope will empower students to achieve in school. The focal participants in the study represented not only students from historically marginalized segments of the population (immigrant, minority, and low-income), but students who had had chronic records of academic underperformance. In two of the four cases, these students were repeating the Introduction to Biology course because of a previous failing grade in the course, or failed to pass the state-issued standardized science exam (a requirement of graduation in the Commonwealth of Massachusetts). As such, these students arguably experienced schooling from marginalized dispositions, where classroom technology use represented an extension of the same institutionalizing forces that relegated them to the lower rungs of academia. Data from this study suggests that students predominantly used technology for multiple-choice assessments, or in ways that mimicked and supported the hegemony of testing activities, such as practice tests, note-taking exercises, and formative assessments that focused on the recitation of facts. Indeed, the there were many similarities between the objectives, mediating artifacts, rules of behavior, and division of labor of these activities and the high-stakes testing environments students typically experienced as a part of schooling. These activities made up the large majority of students' time using technology, and provided arguably few roles for the students to occupy that might have empowered them to transcend the marginalized positions ascribed to them as a result of their poor academic performance, anti-normal social behaviors, and underprivileged cultural values.

A second tension that emerged from the data centered on the way students seemingly undermined the purpose of activities (a form of resistance to schooling), and the use of technology to fulfill course requirements. Over the course of the year, but more so during the first semester, the teacher did in fact assign technology activities that were intended to incorporate elements of student-centered learning and constructivist pedagogy. These activities by and large included creating multimedia presentations, conducting virtual lab experiments, and researching information online. Within these activities, the roles provided for students to become legitimate participants in the knowledge creation of the classroom community extended beyond their traditional academic roles, leaving room for autonomous activity, agency in demonstrating one's knowledge of the subject matter, and the exploration of various digital media that students found of personal interest. Yet, despite a persistent level of encouragement from the teacher, and a far greater freedom to "show what they know" through the use of technology, many of these opportunities were not taken up by the students, who rather, more often than not, undermined the instructor's intent in providing them with autonomy and logistic authority, and further, failed to complete assignments. When such assignments *were* completed, they often demonstrated a lack of resourcefulness in incorporating multiple funds of knowledge, and adhered to the bare minimum of requirements. In other cases, computer technology was eschewed altogether, and students turned in hand written, or hand crafted artifacts that, in some cases, exhibited a greater level of effort and care in their creation than artifacts that were created using technology.

In spite of these findings, the seeming lack of creativity or agency in students' technology-mediated artifacts, or even the lack of self-regulated learning the students took advantage of technology to engage in (e.g., not taking practice tests that included the exact same questions used on actual assessments), did not appear to indicate that students did not possess the fundamental level of technology literacy skills needed to achieve greater academic possibilities. Rather, the focal students exhibited adept use of technologies, especially their cell phones, to achieve both their temporary personal goals (often related to "hanging out" [Ito et al., 2010]), as well as their illegitimate academic ones (e.g., cheating, sharing answers, or using unsanctioned resources to find information). These observed activities resemble "hidden literacies" (Ives, 2011) that are often un-valued, and hence go unseen, in the context of schooling.

Examples of these literacies included using online discussion boards to find information related to the Minecraft video game, using social media (especially Twitter) to create peer networks and participate in peer culture, and "hacking" together disparate (but compatible) technologies to send text messages over the school's network. What was most remarkable about these examples, was that each required the participating student(s) to consciously circumvent school rules, and in some cases, its technology infrastructure. To accomplish this, students had to improvise new ways of achieving their temporary goals that avoided detection from authorities,

including using unsanctioned features of their laptops, such as the "spaces" feature, and the AirDrop file sharing application.

The privileging of certain technology uses in the classroom (taking lecture notes and electronic assessments, podcasting, and digital poster making) arguably contributed to students' sense of "what counts" as learning, and what counted as legitimate course work. Students dismissed opportunities to hold epistemic authority over the content they studied (even during creative media projects), and demonstrated an approach to learning that suggested accomplishing teacher-centered goals, or getting the "right" answer, were the most important outcomes of academic participation.

Students used their occasional logistic authority to avoid engaging with course content (chiefly by "Googling" answers and copying each other's work) and to socialize, rather than planning out effective ways of tackling group assignments or distributing labor amongst themselves to accomplish assignment goals more quickly. These behaviors contributed strongly to the teacher's sense that he needed to reclaim control over his students' learning. He therefore implemented measures that enabled him to manage their engagement with content and on-task behavior more closely. These measures came in the form of increased lecturing, the implementation of more worksheet-guided online activities, and the revocation of both epistemic and logistic authority over their classroom time.

When I asked the teacher for his thoughts on what the class was able to accomplish as a group at the end of the year, he raised the issue of authority and control, lamenting that he did not feel he could sustain these "open-ended" assignments without greater instructional support to both keep students on task, and attend to the range of students' individual learning needs.

I believe these findings speak loudly to issues surrounding not only relationships of power and the hegemony of privileged "ways of knowing" in traditional schooling environments, but of chronic epistemic tensions between institutions', teachers', and disenfranchised students' objectives for school-based learning activities. Regarding the frequency and purpose of technology use, these tensions have the potential to yield outcomes that overlook students' technology backgrounds, and institutionalize the very tools educators hope will empower students for engagement in productive academic and work lives.

Conclusions and Implications

One-to-one computing has gained an astonishing amount of popularity in education in recent years, coinciding with the advent of tablet-based devices and data-driven applications that claim the ability to "personalize learning" at any level. While many have criticized deterministic, "panacea" approaches to educational technology, the wave of one-to-one computing continues to swell. Though often well meaning in focus and in scope, efforts to technologize education over the past two decades have paradoxically contributed to educational inequities that have left historically marginalized populations of students "stuck in the shallow end" (Margolis, 2008). The preliminary findings of this study suggest that without careful consideration, schools risk institutionalizing the very technologies they purport to empower today's youth.

While access to computing and Internet tools, and time to rehearse the various skills needed to utilize those tools, are fundamental aspects to the development of technology literacy, the types of activities students perform with technology, from drill and practice, to research and information analysis, to multimedia production, have an undeniable effect on the technology skills they develop in the context of schooling, and arguably, on the content material that they learn. As such, some have argued that digital education inequity points to tensions that might inhibit the frequency of student technology use, including ineffective (or absent) teacher training and professional development or the lack of instructional support (Belland, 2009). However, many studies leave unaccounted the complexities of resistance in schools and symbolic capital that permeate many communities where educational inequities persist (Anyon, 1980). This study reports on the teaching practices and learning experiences of a learning community where technology integration strategies and access to rich technology learning tools are in long supply. That students would continue to resist schooling in the face of opportunities for enhanced creative expression, access to information, and even to helpful assessment resources, implies a "divide" not in the types of technology literacy skills students acquire, but possibly in the ways schooling legitimizes certain types of learning, and the roles students are allowed to occupy in their learning experiences. This further suggests that schools may need to reconsider how they scaffold students into roles of autonomy, and how they can connect curriculum and pedagogy to students' non-academic technology practices.

References

Anthony, A.B., & Clark, L.M., (2011). Examining dilemmas of practice associated with the integration of technology into mathematics classrooms serving urban students. *Urban Education*, *46*, 1300-1331.

Anyon, J. (1980). Social class and the hidden curriculum of work. *Journal of Education, 162.* 67-92.

Aronowitz, S. and Giroux, H. (1993). Reproduction and resistance in radical theories of schooling. In S. Aronowitz and H. Giroux (Eds.), *Education still under siege* (pp. 65-109). London: Routledge.

Barab, S., Barnett, M., Yamagata-Lynch, L., Squire, K., & Keating, T. (2002). Using activity theory to understand the systemic tensions characterizing a technology-rich introductory astronomy course. *Mind, Culture, and Activity, 9(2)*.

Belland, B. (2009). Using the theory of habitus to move beyond the study of barriers to technology integration. *Computers & Education, 52*(2), 353-364. Retrieved from http://www.sciencedirect.com/science/article/pii/S0360131508001401

Bourdieu, P. (1977), *Outline of a theory of practice*. London: Cambridge University Press.

Corbin, J. M., Strauss, A. L., & Strauss, A. L. (2008).*Basics of qualitative research: Techniques and procedures for developing grounded theory*. Los Angeles, Calif: Sage Publications.

Darling-Hammond, L. (2007). The flat earth and education: How America's commitment to equity will determine our future. *Educational Researcher, 36,* 318-334.

Dimitriadis, G., & Kamberelis, G. (2006). *Theory for education*. New York: Routledge.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Retrieved from http://lchc.ucsd.edu/MCA/Paper/Engestrom/expanding/toc.htm

Giddens, A. (1985). *The constitution of society: Outline of the theory of structuration*. Cambridge: Polity Press.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Pub. Co.

Gutierrez, K., Rymes B., and Larson, J. (1995). Script, counterscript, and underlife in the classroom: James Brown versus "Brown v. Board of Education." *Harvard Educational Review*, 65, 3, 445-71.

Harris, M.J. (2010). Impactful student learning outcomes on one-to-one student laptop programs in low socioeconomic schools. Doctoral dissertation, San Francisco State University, December, 2010.

Hohlfeld, T. N., Ritzhaupt, A. D., Barron, A. E., & Kemker, K. (2008). Examining the digital divide in k-12 public schools: Four-year trends for supporting ICT literacy in Florida. *Computers & Education*, *51*(4), 1648-1663.

Ito, M. (2010). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. Cambridge, MA: MIT Press.

Top of Form

Ito, M., Gutie´rrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., Watkins, S. C. (2013). *Connected learning: an agenda for research and design*. Digital Media and Learning Research Hub.

Ives, D. (2011). Spotting foolbirds: literacies hiding in plain sight in an urban English language arts classroom. *Journal of Literacy Research, 43(3),* 250-274.Bottom of Form

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lemke, C. (2010). Innovation through technology. In J.A. Bellanca and R.S. Brandt (Eds.) *21st Century Skills: Rethinking How Students Learn* (243-72). Bloomington, IN: Solution Tree.

Margolis, J. (2008). *Stuck in the shallow end: Education, race, and computing*. Cambridge, MA: MIT Press.

McDermott, R. (1993). The acquisition of a child by a learning disability. In S. Chaiklin and J. Lave (Eds.), *Understanding practice*. (pp. 269-305). New York: Cambridge University Press.

McDermott, R., & Varenne, H. (1995). Culture as disability. *Anthropology & Education Quarterly*, 26, 324.

McDermott, R. and Varenne, H. (1996). Culture, development, disability. In R. Jessor, A. Colby, and R. Shweder (Eds.), *Ethnography and human development: Context and meaning in social inquir*y (pp.101-126). Chicago: The University of Chicago Press.

Mouza, C. (2004). Learning with laptops: Implementation and outcomes in an urban under-privileged school. *Journal of Research on Technology in Education*, *40*(4), 447-472. *International Society For Technology In Eduction*. Retrieved from http://www.mrgibbs.com/tu/research/articles/mouza_laptops.pdf

Sims, C. (2013). From differentiated use to differentiating practices: negotiating legitimate participation and the production of privileged identities. *Information, Communication & Society*, (September), 1-13. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/1369118X.2013.808363

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge, Cambridgeshire: Cambridge University Press.

Subramony, D.P. (2007). Understanding the complex dimensions of the digital divide: Lessons learned in the Alaskan arctic. *The Journal of Negro Education, 76(1),* 57-67.

Wang, X. & Ching, C. (2003). Social construction of computer experience in a first-grade classroom: social processes and mediating artifacts. *Early Education & Development*, *14*(3), 335-362.

Warschauer, M., Grant, D., Del Real, G., & Rousseau, M. (2004). Promoting academic literacy with technology: successful laptop programs in K-12 schools. *System*, *32*(4), 525-537

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education*, *34*(1), 179-225. SAGE Publications. Retrieved from http://rre.sagepub.com/cgi/doi/10.3102/0091732X09349791

Windschitl, M., & Sahl, K. (2002). Tracing teachers' use of technology in a laptop computer school: The interplay of teacher beliefs, social dynamics, and institutional culture. *American Educational Research Journal*, *39*(1), 165-205.

Zhang, W. (2010). Technical capital and participatory inequality in edeliberation. *Information, Communication & Society*, 13(7), 1019–1039.

Capturing Personal and Social Science: Technology for Integrating the Building Blocks of Disposition

Tamara Clegg, Elizabeth Bonsignore, June Ahn, Jason Yip, Daniel Pauw, Michael Gubbels, Becky Lewittes, and Emily Rhodes
{tclegg, ebonsign, ahnjune, jasonyip, dpauw, mgubbels, charley, eerhodes}@umd.edu
University of Maryland, 2117 Hornbake South Wing, College Park, MD 20742

Abstract: The development of a scientific disposition opens opportunities for youth to see science as relevant in their daily lives. Four building blocks promote disposition development: gaining *competence*, sparking *curiosity*, belonging and contributing to one's *communities*, and bridging *personal connections*. In this paper, we explore the role technology played in supporting unified disposition development by detailing the experiences of two focal learners. We found that technology can act as a boundary object that supports connections across the building blocks in integrated ways, and we include challenges and implications for design.

Introduction

To advance our nation's goal of "science for all learners" (Rutherford & Ahlgren, 1991), researchers note that it is increasingly important to help young learners develop scientific dispositions (e.g., Borda, 2007). The development of a scientific disposition opens opportunities for learners to explore potential roles that they can play in science, whether they choose to pursue scientific careers or to use science practically in their everyday lives (Barton, 1998). We define disposition as values of, ideas about, and ways of participating in a discipline that come frequently, consciously, and voluntarily (e.g,. Gresalfi, 2009). We focus on scientific inquiry as a disposition that will help learners explore their worlds, and we analyze disposition through the lens of scientific inquiry practices. To support learners' efforts to develop scientific dispositions, we must devise ways to help them move beyond abstract facts and phenomena, to extend their classroom experiences beyond the bounds of school (Bereiter, 1995). Specifically, we must enable learners to *scientize* their daily activities, which involves helping learners see the world through scientific lenses, and to integrate this vision in practical applications across the contexts of their everyday lives (Clegg & Kolodner, 2014).

To that end, we have developed *life-relevant learning environments* (LRL) that include programs and technologies designed to help learners to engage in science in the context of pursuing their own personally meaningful goals and to explore potential roles they can play in science. Kitchen Chemistry (KC) is one such LRL program, and is the contextual focus of this paper. KC is an out of school program in which learners engage in scientific inquiry through cooking. Technology has been widely used to support learners' scientific inquiry experiences and understanding (e.g., Barab et al., 2010). Additionally, the ubiquity and culture of technology in today's society, especially among youth (Madden, Lenhart, & Duggan, 2013) suggests that technology holds powerful potential for supporting personal and social applications of inquiry among youth. While technology has been effective at promoting the development of specific aspects of disposition individually, less is known about how technology might promote learners' unified experiences across these building blocks of disposition development. Particularly, we ask: 1) what role can technology play in supporting learners' integrated development across the building blocks of disposition? and 2) how can we design technology to support learners' unified development across the building blocks of disposition?

Background

Based on prior disposition research, we have identified four building blocks that promote learners' disposition development (Clegg & Kolodner, 2014): 1) procedural and conceptual understanding support learners' efforts to develop the *competence* needed to engage in scientific inquiry; 2) interest helps learners develop a *curiosity* about the world – a desire to learn more; 3) social interactions promote learners' engagement in *communities* of individuals who share similar interests as well as communities to which they can make contributions; and 4) personal connections help learners develop personal values for scientific inquiry and reasoning and a *commitment* to engaging in scientific inquiry. While there is a wealth of literature on promoting these building blocks individually, less is known about how we should integrate them into more unified scientizing experiences, and how learners' dispositions develop through more unified experiences. Our previous work in the KC learning environment has shown that LRL environments can be places where the building blocks come together to promote learners' scientific disposition development (Clegg & Kolodner, 2014). We now aim to understand the ways in which we can draw upon the features and affordances of technology to promote learners' unified experiences across the building blocks of disposition development.

Given the ubiquity of technology in children's lives, they are engaged in a multiplicity of personally relevant, technology-based experiences, such as gaming (e.g., Squire & Jenkins, 2003) and social media

participation (boyd, 2009). Through these deeply personal and interest-driven experiences, learners participate increasingly in affinity groups (Gee, 2005), which foreground common interests over membership in more formal structures. Key to unifying learners' development in the building blocks of disposition is then helping them to connect their life-world subcultures to the subcultures of science (Aikenhead, 1996). One potentially transformative factor in helping learners make these connections is the creation and management of *boundary objects* (Star & Greisemer, 1989; Wenger, 2000), or artifacts, tools, and processes that help people from different communities collaborate in meaningful and productive ways (Star & Greisemer, 1989; Wenger, 2000).

Our design-based research efforts (Collins, Joseph, & Bielaczyc, 2004) over the past few years suggest that technology holds potential to be an effective boundary object for bringing disposition building blocks together. Our initial efforts in life-relevant learning focused on connecting learners' scientific inquiry practices (i.e., procedural and conceptual understanding) to their personal interests (Clegg et al., 2012). We found that certain affordances of mobile technologies (e.g., storytelling) can help learners mediate between their interests, personal connections, and scientific engagement to promote unique approaches to and expressions of science. Next, we used a social media tool specifically designed to support learners' collaborative scientific inquiry (Clegg et al., 2013). We found that technology can help diverse learners recognize common interests and build on the contributions of others, even in tense social environments, to facilitate productive social collaborations in science (Clegg et al., 2013). While this work suggests the potential of technology as a boundary object to bring together the building blocks of disposition, few studies exist that examine how technology can be designed and used to promote an integrated experience across the four building blocks of disposition.

Our Approach: Life-Relevant Learning

While previous work looks at bringing subsets of the disposition building blocks together, the question remains: How do we design technology to integrate all four into a unified experience? Existing technologies have attempted this integration, but tend to foreground procedural and conceptual understanding. This makes it less likely learners will *want* to pick up the tool on their own in experiences that are important to them. Our approach has been to foreground the personal, using it as a means of promoting further learning. We have done this through the design of (1) LRL technologies and (2) LRL programs.

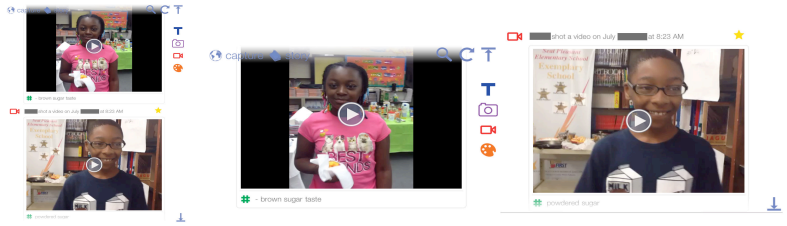


Figure 1. Timeline view in ScienceKit.

One component of our work is designing a LRL technology - *ScienceKit (SK)* (Figure 1) - that integrates the affordances we have found useful for unifying learners' experiences across the building blocks of disposition. SK is a mobile and social app that allows learners to capture and share snippets of daily life – similar to social media tools such as Instagram – but frames these sharing practices through a lens of scientific inquiry. SK is designed to scaffold learners' scientific inquiry in their daily lives by enabling them to create micro-contributions to the inquiry process. In SK, learners create entries with their choice of photos, videos, text, or drawing. With these media, they develop questions, observations, experiment sequences, cause and effect claims, or "just because" entries. These entries are shared on a sequential, public timeline for everyone to view. Learners can then "star" contributions as a means of favoring them. We used SK in the KC implementation presented in this paper. In the initial sessions of KC, learners engage in *semi-structured activities,* becoming familiar with processes in measurement, data collection, and technology usage in the context of cooking experiments aimed to answer scientific questions (e.g., What do eggs do in brownies?). On *Choice Days,* learners are given opportunities to use what they learned to develop questions, hypotheses, experimental procedures, and data collection techniques for their own food investigation. They also make decisions about recipe modifications, controlling variables, data collection, and interpretation of their findings. We designed these experiences in KC to support learners' scientific practice through interest-driven experiences with peers and adults to support learners' development across the building blocks of disposition development.

Methods

For this study, we employed the methods and standards of a comparative case study (Yin, 2003). The case is a single 1-week summer camp implementation of KC. In this exploratory study, we focus on the role of technology in supporting learners' integrated experiences across the building blocks of disposition development

as we highlight their experiences in KC. We focus this case study on two learners: DeMarco and Allen (all names are pseudonyms). We selected these two children for this study as a comparison between how technology can support scientific inquiry and disposition development in sociable (DeMarco) and reticent (Allen) learners.

Context and Data Collection

For this implementation of KC, we met as an out-of-school, summer camp program for four consecutive half-days (Monday – Thursday, 4.5 hours per day) in a lower socioeconomic status public school in the Washington DC metro area. Seven learners (9 to 11 years old) from the school participated in the program. Our KC implementation was comprised of eight adult facilitators, one of whom focused primarily on technical support. The first two sessions were semi-structured days and the last two were Choice Days. On Day 1, learners observed brownies made with different amounts of eggs and did an experiment with eggs, oil, and water to understand how eggs work in brownies. Day 2 involved a cookie experiment to test and explore the roles of different leaveners. On Days 3 and 4, learners chose new dishes to perfect and worked on their Choice Day investigations with facilitators. Each day we collected video recordings of all activities and discussions. Additionally, to understand learners' identity development as they participated in the program, learners created short *personal* reflection videos in SK at the end of Days 1 - 3 in which they responded to the prompt: "*Today, I was more like a … Chef, Investigator, Scientist, A combination (tell us which combination you felt like), or something else (tell us what or who). Then, tell us what you did to make you feel like that."* We conducted interviews with each learner on the third day of the program, focusing on their use of SK, their experiences in science, and their experiences in KC. Facilitators also recorded post-observational field notes of their experiences each day in KC. Lastly, we collected analytics (e.g., time stamps, account logins, SK posts) as participants posted contributions to the SK app.

Data Analysis

Our data analysis process included two phases. In the first phase, we analyzed data types individually. For each learner's personal reflections, we transcribed and conducted open coding (Corbin & Strauss, 2008) to identify themes for each learner based on the identity-related information they provided. Based on this analysis, we selected two focal learners with the most drastic differences in participation styles: very sociable to very shy. This comparison is important because participation styles play a significant role in learners' disposition and identity development (Clegg & Kolodner, 2014) and we want our tools and programs to support a diverse range of learners. Next, three authors coded the SK posts of the two focal learners, describing each media file and coding entries for aspects of learners' scientific practice based on Chinn and Malhotra's (2002) inquiry framework. We also coded for themes related to the types of interest-based, social, and personally meaningful experiences learners were having with the technology (Clegg et al., 2012). Finally, four other researchers built contextualized stories of each learner's experience in the program, using a combination of observation data (e.g., post observation field notes, video transcriptions). While SK entries represented snapshots of specific experiences learners had, contextualized stories provided a broader perspective of learners' daily experiences. The second phase of our analysis consisted of a collaborative analysis session in which all three data types were integrated. All researchers gathered and printed out each set of data (with media files from learners' SK entries and photos from the contextualized stories). We lined each learner's data sources up side-by-side sequentially to represent that learner's experiences with SK. The visual nature of our data enabled us to have an axial coding session in which researchers used sticky notes to create analytic memos that were grouped into major themes for each learner. We followed with selective coding (Corbin & Strauss, 2008), to identify themes across cases.

Findings

Here, we present the cases of each focal learner, DeMarco and Allen, organized by the themes underlying our research questions.

DeMarco: A Social Entry Into Science

DeMarco was a rising 5th grader who entered KC with a calm demeanor but took a social approach to scientific inquiry throughout the session. DeMarco's initial participation in KC involved using SK to take photos and videos of other learners in the program. During almost all whole group conversations, DeMarco recorded conversations with SK, moving the camera to record the speaker. DeMarco actively participated in these discussions, contributing thoughts and questions as he recorded. He began to ask questions in KC at breakfast on the first day. As the learners ate Apple Jacks™ cereal, a facilitator, Naomi, wondered if the cereal actually contained apples. She referred to the ingredients label on the container and was surprised to see that they did contain apples. Seeing DeMarco's interest in the question, Naomi encouraged him to create a question in SK. DeMarco then shifted from recording his friends at breakfast, to creating a video about the question they had, showing the ingredients label to answer the question. Later, during breakfast, DeMarco repeated this process of creating question videos when he heard another learner ask a similar question about milk. In his video, DeMarco

displayed and read the ingredients label on the milk carton, then concluded that milk does not have sugar. He continued to ask questions about ingredients that he was interested in during the following days' activities.

Semi-Structured Sessions: Scientific Reporting

DeMarco continued to take this reporter role in the initial sessions of KC. In this role, he documented and narrated the types of experiences they were having in KC, often interviewing group members about what they were doing. DeMarco's reporting became more scientific on Day 1 as John, the facilitator working with DeMarco's group, was careful to model the types of observations to make and ways to use SK to record observations. On his own, DeMarco took photos of their experiment variations, paying close attention to capture details about the variations in his photos (e.g., variation in ingredient quantities, before and after photos, etc.). As he made observations about their experiments, he also began to make predictions and claims based on those observations about the differences in the brownie samples, and how eggs work in brownies. DeMarco's reporter role also involved taking photos and short video clips that documented their individual experiment procedures, recording measurements of their cookie heights, and documenting their procedures for measuring.

Personal Science: Fried Chicken Experience

DeMarco worked with facilitator, August, to plan his Choice Day experiment. This was a very personal and scientific experience for DeMarco, and it involved significant social motivations for him in and out of KC. DeMarco's idea arose because he wanted to make the fried chicken his mom cooks at home. He expressed that she never let him in the kitchen and he thought she would be proud of him for making a complex recipe. Facilitators noted that it "made his day" to be able to make fried chicken. However, due to safety concerns, DeMarco's project also was a source of much consternation for the facilitators. These concerns were a great source of humor and helped develop rapport between DeMarco, facilitators, and other community members (like the school lunch lady – Ms. Carlitta). DeMarco's pride in this project was illustrated as he posted sketches of fried chicken in SK (e.g., drawing a chicken leg with a heart around it) during planning on Day 2.

As DeMarco engaged in his Choice Day project, two primary ways he engaged in scientific inquiry were: 1) experiment planning and 2) observing and analyzing experiment results. DeMarco worked with facilitators John and August to plan his Choice Day experiment. John suggested they fry the chicken in different oils to learn about the differences. The facilitators reported that he warmed up to the idea when he saw that they (the facilitators) were excited about it. He even brought his mom's olive oil on the first Choice Day (Day 3) to use in the experiment. They planned a 2 x 2 experiment that started with testing different chicken breaders (i.e., flour versus flour + egg wash) and oils (i.e., olive oil vs. canola oil). On Day 4, the group cooked each breader variation in different oils. As the group conducted different experiments, they had DeMarco observe and analyze the results in different ways, such as blind taste tests with the results with DeMarco describing the results as he tasted them. Facilitators reported that DeMarco was excited to see observable differences in the color and texture chicken fried in olive oil versus canola oil (regardless of breader variation). As DeMarco described the results of each variation of chicken, August created two graphs to visualize his results.

Challenges with Scientific Inquiry

Although DeMarco engaged in these scientific practices during Choice Day, facilitators observed that he still needed significant scaffolding to think of his fried chicken in a scientific way. The facilitators found it more challenging than previous days to get DeMarco to think scientifically in the context of such a personally meaningful endeavor. They had difficulty helping him move from opinion to more descriptive observation (e.g., moistness and crispiness as observable outcomes versus "it tastes good"). Additionally, August thought DeMarco was "a little lost" during the graphing exercise and more generally, less interested when they talked about the science. DeMarco's SK posts also demonstrate some of the challenges he faced engaging in scientific inquiry during Choice Day activities. Of a total of 107 entries that DeMarco created throughout the KC summer camp, only 16 of them were created on Days 3 and 4 (Choice Days). The Choice Day entries were primarily *recordings of* DeMarco by others and not entries *created by* him. Personal and social photos were taken of DeMarco, and science videos were posted of facilitators interviewing him about their experiment procedures.

DeMarco's Personal Positioning in Science

Although the facilitators noticed a decrease in DeMarco's interest in scientific inquiry during their Choice Day project, they observed a resurgence of energy and engagement during the children's oral presentations. In preparation for showing their parents, families, and the school community about their Choice Day experiences and products, DeMarco set up his presentation behind the librarian's desk. He used both the chicken he made and the graphs generated during their reflections to present his results to the community. He described to his audience the experiment they did and explained the results represented on the graphs with competence and gusto. Based on their data, he told his audience he would suggest to his mom a different oil to use in the future. In field notes, the facilitators observed the pride, confidence, and authority with which DeMarco presented his

results. One facilitator noted, "He totally owned his graphs as he explained his results to others." As DeMarco carried out his experiment, he received recognition from others in KC, his own family members, and the larger school community. Another facilitator noted that the other KC children often viewed a photo DeMarco posted to SK in which he posed confidently with his chicken, showing the gloves he wore as he prepared the fried chicken. Facilitators also noted that DeMarco "just beamed" as he presented his work to the broader community and received accolades not only from his siblings but also from others' family members and the school staff.

Role of Technology in Promoting DeMarco's Disposition Building Blocks

DeMarco's technology use matched his experience in the physical environment. First, he used SK to document his *social* experiences in KC using videos. With modeling and prompting from facilitators, he began to use SK to express the *interest* he had in new questions and the observations (*procedural and conceptual understanding*) he made using SK's integrated multimedia. While these observations were scientific in nature, many simultaneously reflected the social and playful experiences DeMarco had during KC. He referred back to these entries later in KC discussions as well as in his personal reflections as evidence for his participation as a chef and scientist. During Choice Day, he made contributions that expressed his *interest and personal connections to* making chicken (drawing, photos). He posted photos and videos that showed his engagement in the process. Although DeMarco also had scientific contributions during these highly personally experiences, his use of SK dropped significantly. However, the contributions he did create were widely viewed (on the timeline) receiving a total of 14 favorites, helping DeMarco to be recognized socially for his contributions.

Allen: Seeing Science Through the Lens of ScienceKit

In contrast to the social antics of DeMarco, Allen's overall case narrative and associated vignettes reflect the ways in which SK supported a painfully shy learner's efforts to develop his building blocks of scientific disposition. Allen often acted nervous and quiet, during whole group discussions and in smaller cooking groups. He was hesitant to talk unless explicitly prompted, even in one-on-one situations. When facilitators addressed him in small group settings, he wavered, often putting his hand to his head or mouth and delaying a response. At times, he did not speak at all, and would duck behind his iPad™ screen, as if waiting for the group's attention to transition from him. Allen also appeared embarrassed about being "on camera," particularly during his personal reflection videos or when other KC learners tried to record him. In their small groups, facilitators often observed Allen distancing himself from his group. Because of the ways in which he held back during small group activities, facilitators were concerned that he was not engaging in scientific inquiry. One facilitator's field notes captured this sentiment with her hope that Allen would soon "put one foot in the experiment instead of none."

The First Day: Connecting with Others through Minecraft

Although Allen was timid in his interactions with others, he appeared more focused and adventurous in his interactions with SK. During our initial review of facilitator field notes and learners' personal reflections, it appeared on the surface that Allen did not evolve as much as we had hoped in terms of scientific disposition and inquiry practices. His silence and physical distance during small group interactions made it difficult to detect any overt progress. However, Allen's SK data, corroborated by vignettes from the session videos, indicates that Allen did advance across several disposition building blocks, by participating through the lens of ScienceKit and the iPad™. Like most learners in this group, Allen immediately began by using SK tools to make sketches. Allen connected his initial forays with SK to his personal interests by drawing scenes and images about the game Minecraft™. One result of SK's social media design was that Allen's drawings were seen in the public SK timeline. One facilitator, Naomi noticed Allen's drawings of "creepers" (characters in Minecraft™) and engaged him in a discussion about it. These small moments were important ways that Allen developed rapport and connection to the KC program, despite his outwardly shy persona.

Through the Eyes of Allen: ScienceKit as an Expressive Tool

As the week progressed – in contrast to his aversion to being recorded by others – Allen would sometimes talk, sing, or dance to his iPad™, in his own personal "selfie" SK videos. In one vignette, he made a short video of himself singing, "Here we're cooking." This example was corroborated both in SK data analysis and in our review of the session camera video that captured him standing apart from, and unnoticed by others working in the environment at the time. This was just one instance of several playful and focused connections he made to share his personally meaningful experiences through SK.

Perhaps the most compelling evidence of Allen's efforts to connect personally to scientific inquiry were the sketches he made of his Choice Day experiment, which involved making s'mores, a marshmallow, chocolate, and graham cracker sandwich. In the beginning, Allen focused most of his energy on connecting personally to this project. Daniel, a facilitator, noted that because Allen was less vocal about his ideas than the other learners, his drawings often functioned as a communication tool throughout the planning and execution of his s'mores' investigation. Many of Allen's SK entries on s'mores were sketches, especially those completed as

he and Daniel planned his experiment. For example, he imagined his s'more as a mythical Minecraft character (the "herobrine"), and included the written description, "they will look like a herobrine skin, lol, I am making s'mores." This entry proved to be one way in which he also tried to connect socially with Tonya, a facilitator, as she asked him what questions he wanted to answer with his experiment. In the video vignette, Allen located this entry in the SK public timeline and showed it to Tonya, after she had asked what s'more shapes he wanted to evaluate. He explained that his s'more was a "herobrine s'more," and he wanted to add eyes and a mouth. Of note, the square structure of a s'more echoes the pixelated block that is the base building unit in Minecraft. This interaction not only offers evidence of attempts by Allen to connect his personal interests to his scientific inquiry efforts, but also highlights his attempts to communicate this connection to others.

Allen's Scientific Dispositions through the Lens of SK

Our session videos and SK data indicated that Allen was actively involved in documenting scientific procedures and making observations. During the semi-structured activities, he would take short videos of experimental steps (e.g., shaking water bottles), sometimes recording the procedure as he was taking part in it. He also took a photo of a brownie through a magnifying glass to show a closer view of its surface irregularities. Although we did not observe Allen recording measurements (e.g., the diameter or area of the s'mores), he made several SK entries of the other KC learners, with accompanying text. Thus, SK's image and video capture features enabled Allen to develop *procedural and content knowledge* aspects of his scientific disposition. Allen also used SK to make predictions or claims about experimental outcomes, and to pose questions about his observations. For example, when prompted, he contemplated reasons why different batches of brownies exhibited distinct textures. Of note, these SK entries contrast with his first day working in a KC group, during which time he did not verbally offer any explicit hypotheses or claims. In addition to scientific observations that Allen made during group experiments, he documented experimental results through SK. Sometimes he would capture one variation at a time in a separate entry until he had captured them all. At other times he would take photos of all variations or several at one time, looking across them. In one entry, he used SK to sketch "good" versus "bad" cookies by highlighting brown, burned edges for the "bad" cookies. Allen's active capturing in this pattern demonstrated his being attuned to the experiment at hand and comparing across variations.

Recasting Minecraft S'mores as Engineering

As noted earlier, Allen chose to make s'mores for his Choice Day Project. As his s'mores planning developed, he made SK entries that concretized his ideas, and helped him imagine variations that were both playful and engineering or design-based. For example, he imagined and sketched s'mores as a house, a jail, and a different Minecraft character. His SK sketches became increasingly more engineering and design-oriented after his experiment. For example, one SK entry reflects a workflow how-to reference for assembling s'mores: "cracker → cracker + chocolate → cracker + chocoloate + marshmallow" (note the sequential arrows as visual cues). He also continued to practice other dispositions in his Choice Day project. For example, he took photos comparing the different s'mores ingredients (e.g., baker's chocolate versus milk chocolate bars) and evaluating his finished s'mores. From an overtly social perspective, however, Allen still hesitated when carrying out tasks and making decisions, despite the fact that the s'mores investigation was his Choice Day design.

Finding a Niche in KC

Over the course of the week, Allen earned a social reputation and identification as an experienced iPad™ user. Facilitators would point toward Allen's iPad™ use as an example of best practices (e.g., switching his iPad™ camera lens). Fellow learners sometimes asked him for help, or followed his (mostly silent) lead. In contrast with his engagement in the physical environment and activities, during which he hesitated or made mistakes because he did not inform facilitators or ask timely questions, he was meticulous about his SK entries. In one video vignette, Allen's longest unprompted discussion with facilitators (~3 minutes) occurred because he had a question about how SK loaded media onto its timeline. He was frustrated that some of his work might be lost due to delays in loading media, or server crashes. In response to Allen's thoughtful, designer-oriented concerns, Daniel suggested that he record his question as a design idea for future iterations of SK, and Tonya recommended that he confer with Michael, the SK developer who could attend to his ideas more deeply.

Role of Technology in Promoting Allen's Disposition Building Blocks

While facilitators had difficulty seeing Allen's scientific disposition develop in person, our analysis of his SK entries and video data revealed a different story. We observed evidence that Allen used SK to begin to make personal connections to science-based KC activities, increase his procedural awareness of scientific inquiry (e.g., making observations), and to initiate social connections in ways that he could not have done without a technology support like SK. Allen connected with SK as he first used it to relate to his *personal connections* (drawing and Minecraft™). He continued to develop this relationship as he had expressive moments with the iPad™ not observed by others. He used it to connect to *scientific inquiry practices* as he meticulously

documented his observations and experiment procedures using SK's multimedia capture features (photos, videos and sketches). Not only did Allen document scientific elements when prompted, he also documented personally meaningful aspects of his scientific experiences. Furthermore, he described and articulated his *interest* in making s'mores through drawing, and used SK to communicate with facilitators about his ideas. In this way, SK served as a mediator for Allen's social interactions.

Discussion: Boundary Encounters and Objects

Science education literature establishes that science learning is a matter of border crossing between multiple communities (Aikenhead, 1996). In fact, learners transition between sub-cultures of home, school, peers, and media everyday. Their engagement in science learning is a matter of becoming enculturated into the customs and practice of scientific inquiry (Aikenhead 1996). Participating in KC can thus be seen as a boundary encounter for learners – a place where they were confronted with the idea that cooking and science were related and valued activities. KC was an environment where learners were introduced to a form of social and interest driven science that seemed new to them. Allen and DeMarco's initial hesitance to share their own ideas suggests that this was indeed a new form of science engagement and learning for them. The people, processes, and norms in KC were quite different than school (e.g., working with researchers, cooking, coming up with new scientific ideas to explore). Promoting learners' development across the building blocks in this context thus required helping learners connect to a new community socially, personally, and scientifically. It also involved helping them connect their participation in outside communities to their participation in KC in meaningful ways.

Role of Technology in Disposition: ScienceKit as a Boundary Object

Boundary encounters are facilitated through boundary objects. We conjecture that SK served as a boundary object that could potentially enable multiple building blocks to be integrated into one tool for many different types of learners. The diversity of Allen and DeMarco's participation styles shaped different life-world connections and scientific progressions for each of them. DeMarco connected to scientific engagement through first linking it to his social practices and home values, while Allen first connected SK to his interest and personal connections to the gaming community of Minecraft™.

Wenger (2000) discusses three types of boundary objects for connecting diverse communities: processes, discourses, and artifacts. We found that SK served as an artifact within the environment that helped learners connect different aspects of their participation, and also influenced the processes and Discourses of the environment. Learners developed norms of use (processes) for using SK that included taking on the role of reporter to document what was happening *socially* and *scientifically* in the environment, and capturing their own *interests* and *scientific engagement*. These norms also included scrolling through SK to see what others were doing. Learners' continuous browsing through SK enabled DeMarco to be recognized *socially* for his contributions in KC by others early on and later as others continuously viewed his contributions to SK. SK also shaped the language of science (Discourse), helping it to become more *personal* and *social* with multi-media and social media features. As an artifact, SK served as a repository of learners' experiences that helped us to see Allen's previously unnoticed participation and those that DeMarco often referred back to during his personal reflections. The repository represented this new form of social, personal, and interest-driven science by integrating the four building blocks of disposition through its technical features.

Design Implications and Challenges for Disposition Development

Even as an effective boundary object, some aspects of SK presented challenges for learners, educators, and designers. DeMarco's case showed that in boundary spaces, if learners move too far over into one building block (e.g., personal connections), it can be more challenging to motivate their scientific participation. DeMarco's case suggests that boundary objects may be more effective at integrating the disposition building blocks if designed to foster additional connections to broader communities (e.g., family, peer, or media) that are motivating audiences for learners. For example, future iterations of SK could include a means for learners to share self-selected media artifacts outside the local KC community (e.g., with family). Allen's case reveals another boundary space in which learners participated – the space between the virtual and the physical. SK primarily supported Allen's integration of the building blocks in the virtual environment. However, he needed more help with social interests in the physical environment, particularly with respect to social interactions and being recognized. His case illuminates the potential to use data and analytics from applications like SK, to help make educators and others aware of learners' participation. These in-situ data snapshots can strengthen our awareness of the link between learners' virtual and face-to-face participation, and suggest opportunities for authentic, formative evaluation of disposition development.

Conclusion: Contributions of This Work

The work presented in this paper contributes to our understanding of how to promote disposition development with technology. We appropriated common media sharing features of social media in the design of SK. By

allowing children to capture and share their daily experiences, through a lens of inquiry, SK helped learners integrate the disposition building blocks in ways that began to promote learners' comprehensive scientizing experiences. This analysis suggests that viewing technology as a boundary object is helpful for understanding how to better promote such unified scientizing experiences for learners. When we recognize practices, tools, and artifacts as boundary objects we can better design them to promote learning (Wenger, 2000). Therefore, pointing to technology as a boundary object for promoting disposition can foster a useful perspective for designing new technologies (of all types) to promote disposition. This work also highlights challenges that remain in understanding how to best design and use technology to promote disposition development. Specifically, it points to the need for understanding ways to effectively connect learners' participation to broader communities and for using learning analytics as a means to alert educators of learners' progress. More work is needed to understand the role of technologies that serve as boundary objects over longer periods of time and how such technologies can be integrated into learning environments to successfully promote disposition.

References

Aikenhead, G. S. (1996). Science education: Border crossing into the subculture of science. *Studies in Science Education, 27*, 1-52.

Clegg, T., & Kolodner, J. (2014). Scientizing and Cooking: Helping middle-school learners develop scientific dispositions. Submitted to *Science Education Special Themed Collection on The Intersection of the Learning Sciences and Science Learning in Everyday Life*.

Clegg, T., Yip, J. C., Ahn, J., Bonsignore, E., Gubbels, M., Lewittes, B., & Rhodes, E. (2013). When face-to-face fails: Opportunities for social media to foster collaborative learning. In Proceedings, *CSCL 2013*, International Society of the Learning Sciences, 113–120.

Clegg, T., Bonsignore, E., Yip, J., Gelderblom, H., Kuhn, A., Valenstein, T., … Druin, A. (2012). Technology for promoting scientific practice and personal meaning in life-relevant learning. In *Proceedings Interaction Design and Children 2012*, ACM Press, 152–161.

Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2010). Relating narrative, inquiry, and inscriptions: Supporting consequential play. *Journal of Science Ed. and Technology*, *19*(4), 387-407.

Barton, A. C. (1998). Teaching science with homeless children: Pedagogy, representation, and identity. *Journal of Research in Science Teaching*, *35*(4), 379-394.

Bereiter, C. (1995). A dispositional view of transfer. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 21–34). Mahwah, NJ: Erlbaum Associates, Inc.

Borda, E. J. (2007). Applying Gadamer's concept of disposition to science and science education. *Science & Education, 16*(9-10), 1027–1041.

boyd, D. (2007). Why youth (heart) social network sites: The role of networked publics in teenage social life. In D.Buckingham (Ed.), *Youth, identity, and digital media* (pp. 119–142). Cambridge, MA: MIT Press. Press, Cambridge, MA, pp. 119-42.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175–218.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15-42.

Corbin, J., & Strauss, A. (Eds.). (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage.

Gee, J. P. (2005). Semiotic social spaces and affinity spaces: From The Age of Mythology to today's schools. In D. Barton & K. Tusting (Eds.), *Beyond communities of practice: Language, power, and social context* (pp. 214–232). New York: Cambridge University Press.

Gresalfi, M. S. (2009). Taking Up Opportunities to Learn: Constructing Dispositions in Mathematics Classrooms. *Journal of the Learning Sciences, 18*(3), 327–369.

Madden, M., Lenhart, A., & Duggan, M. (2013). Teens and Technology 2013. Pew Internet & American Life Project. Retrieved Oct 21, 2013 from: http://www.pewinternet.org/Reports/2013/Teens-and-Tech.aspx.

Rutherford, F., & Ahlgren, A. (1991). *Science for all Americans*. New York: Oxford University Press.

Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. *Insight*, *3*(1), 5-33.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387-420.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, *7*(2), 225–246.

Yin, R. K. (Ed.). (2003). *Case study research: Design and methods* (Vol. 5). Sage.

Acknowledgments

We would like to thank our KC participants, faculty and parents at the school in which KC was held.

Moving Beyond Case Studies: Using Social Network Analysis to Study Learning as Participation in Communities of Practice

Julia Eberle, Karsten Stegmann, Frank Fischer
Ludwig-Maximilians-University of Munich, Leopoldstr. 13, 80802 Munich, Germany
julia.eberle@psy.lmu.de, karsten.stegmann@psy.lmu.de, frank.fischer@psy.lmu.de

Abstract: We argue that social network analysis is a useful methodology to study learning under the participation metaphor and to extend scholarly knowledge on the process of legitimate peripheral participation in communities of practice. We emphasize that it is necessary to extend current research about the epistemic aspects of participative learning and to take also the social aspects into account; we first review work on legitimate peripheral participation and how research on this topic is currently conducted. Then we describe the theoretical and methodological foundations of social network analysis. Exemplary studies that use social network analysis including studies on legitimate peripheral participation are presented. We conclude that social network analysis can enable us to make inferences about participative learning across specific setting and individual communities. For the future, the challenge is to relate dynamically changing patterns of legitimate peripheral participation to emerging and changing epistemic practices of a community.

Introduction

The perspective of learning as becoming in practice is closely related to what Sfard (1991) termed the participation metaphor of learning. Learning as seen through the lens of this metaphor focuses on the development of individuals' identity by engaging in certain practices and participating that way in a specific social environment. This approach seems especially valuable for research on informal and workplace learning in which learning goals and paths are very individual and can often hardly be defined by outsiders. Learning under the participation metaphor takes into account that the process of learning and the individual learning goals are situated in the learner's social and epistemic environment. However, most research in this direction still focuses on the epistemic dimension of participation, by analyzing practices and artifacts produced by participative learners. The social dimension has only rarely been the focus of research, probably in part because of a lack in research methodologies that would allow for meaningful ways of analysis.

In this paper, we argue that social network analysis (SNA) is a very promising methodology to study learning under the participation metaphor. SNA is not a new methodology to the ISLS community. Yoon (2011), for example, introduced SNA as a visualization tool to foster students' interaction. It also has been used to analyze community building in scientific communities (Hoadley, 2005). However, using SNA in a pure methodological way to operationalize and analyze learning processes and outcomes is still a mostly unexplored field in the Learning Sciences.

To show how SNA can be applied as a research methodology to enhance our understanding of learning as participation, we first review work on the theoretical approach behind the view of learning as participation and the empirical evidence gathered so far. We will then argue how SNA can be applied to study learning under the participation metaphor, followed by two exemplary studies in which this methodology has been applied. In the conclusion, we come back to the question to what extent SNA is a useful methodology to study learning and what possibilities it provides for future research.

Learning as Participation

Research with the perspective of learning as participation mostly focusses on the understanding of real-world learning occurring outside of formal learning situations. It is a situated approach to learning based on the idea of legitimate peripheral participation of a learner in a community of practice which is considered to lead to a transformation of the learner's identity (Lave & Wenger, 1991; Wenger, 1998). Through participation in activities which are specific to the community of practice, a newcomer is confronted with the community's artifacts, and has opportunities to learn the relevant practices, and participates in the community's process of negotiating meaning.

Learning as Legitimate Peripheral Participation in Communities of Practice

One of the central aspects of learning under the participation metaphor is the epistemic and social context in which learning takes place. Lave and Wenger (1991) termed this context "community of practice" (CoP). Capturing the CoP concept adequately is tricky, although (or probably even because) it is widely used and applied by researchers and practitioners in several disciplines and professions. However, Barab and colleagues have worked on a comprehensive set of characteristics that define a CoP and make the concept graspable and

 © ISLS

workable with for Learning Scientists (Barab & Duffy, 2000; Barab, MaKinster, & Scheckler, 2003). They define a CoP by several epistemic and social aspects: shared knowledge, values and beliefs, overlapping histories and mutual interdependence among members, mechanisms for reproduction, a common practice and/or mutual enterprise, opportunities for interactions and participation, meaningful relationships, and respect for diverse perspectives and minority views.

This set of characteristics clearly shows the importance of both epistemic and social aspects that characterize the CoP concept, although most research based on the concept has so far emphasized the epistemic part, especially the shared practices, and more or less neglected the social aspects. However, Wenger, McDermott, and Snyder (2002) have made some further claims on the social structure of a CoP and differentiate between members that belong to a small core group that leads the CoP, a larger group of active members that run the CoP, an often even larger group of peripheral members that only participate irregularly in the CoP, and finally outsiders who might be influenced by the CoP but are not directly involved. Some CoPs also have a coordinator to facilitate the actions among the CoP members. Another way of speaking about the social structure, especially relevant for learning, is the differentiation between newcomers and old-timers who engage in the CoP in different ways. The transition process from being a newcomer to becoming an experienced old-timer in the CoP was described by Lave and Wenger (1991) as legitimate peripheral participation (LPP).

They describe LPP as a natural form of learning which happens all the time, no matter whether there is an intended educational situation or not. The main focus of this perspective on learning is an increase of newcomers' level of participation within a specific CoP. A newcomer usually starts as a mostly passive outsider and becomes a more capable and active member over time, accompanied by a transformation of his or her identity. This transformation includes a change in cognition, communication, and behavior towards the shared practices of a CoP and is based on a CoP-specific, socially shared learning history. These changes cannot merely be found in individual minds but also in the increased participation in CoP practices and tasks. Also, the relations between newcomers and more experienced members develop with the newcomers' engagement in the practices of a CoP practices.

Peripheral participation means an increase of newcomers' involvement, starting with rather passive activities in the beginning like observing how old-timers talk, interpret new situations, react to them, and deal with problems. Newcomers can take over simple and easily manageable tasks that do not require much knowledge about the CoP. Later, when newcomers have learned more, they can take over less peripheral and more complex tasks that require more responsibility and are more important to the CoP. One of the remarkable differences between LPPand learning in deliberately designed educational settings is the "design" of the curriculum. The curriculum in the LPP process is based on the practices and activities that occur in usual situations the CoP faces, but does not follow the steps an old-timer would follow to achieve a CoP goal (Lave & Wenger, 1991). Newcomers are only allowed to contribute to these activities of goal achievement that do not require more than they can accomplish because this would bare the danger for the whole CoP to fail their goal. This gives newcomers opportunities to actively participate in CoP life and get insight into its practices without having the role of a non-involved visitor. Newcomers have access to practices that are typically not explicated, like how individual members of the CoP interact with each other and with outsiders of the CoP..

The second important aspect of this kind of learning is the question of *legitimation* of newcomers' participation. Lave and Wenger (1991) assume that there are more or less strict entry criteria or barriers in each CoP. These entry criteria regulate newcomers' possibilities for learning. Only legitimate newcomers get access to community activities, information and resources and those are important for peripheral participation. Several case studies (for example Back, 2011; Fuller, Hodkinson, Hodkinson, & Unwin, 2005) have illustrated the relation between newcomers' legitimation in the CoP and their access to possibilities for peripheral participation. Members to whom the CoP does not grant access as legitimate members for some reason, cannot become more active participants. They form a special type of peripheral members, called marginalized members. Marginalization can have various reasons, most common are reasons related to the broader context of the group like hierarchies or positions in an organization that prevent that some persons can become members of a particular CoP. Certain characteristics, like gender, age, or ethnicity of marginalized persons can be other reasons for marginalization.

Lave and Wenger (1991) stated that the LPP process is dependent on how the resources that are important for newcomers' learning are structured in the community or how transparent the CoPs artifacts are. There is usually not much explicit teaching involved, but there can be masters, sponsors, and other newcomers who might help structure the learning process and make invisible artifacts a little more transparent. Communication plays therefore an important role for learning in CoPs, not only for transferring information but also to guide the attention of newcomers, engage them and coordinate collaboration. But according to the authors, most important for newcomers' learning are possibilities for participation. It can therefore be assumed that newcomer's exposure time to a CoP will play an important role for the learning process: The more time a newcomer can spend with CoP activities and other members of the CoP, the more opportunities for learning will arise. LPP should not be understood as "mere exposure effect" in which the amount of time spent in a CoP is the

only crucial factor for a newcomer's level of participation, as Lave and Wenger's examples of influencing factors show. They identify a research gap at that point wondering how the knowledge of CoPs can be made transparent to newcomers.

Research on Legitimate Peripheral Participation

Compared to the CoP concept, the LPP concept as the core element of the initial theory has been studied less intensively. However, it was applied to workplace learning and has proven to be a useful approach, for example in workplace learning of nurse teachers (Boyd & Lawley, 2009) and of police officers (Campbell, Verinikina, & Herrington, 2009). Researchers have also investigated, if online environments can be created to build CoPs for professionals in order to support newcomers to professions in which they usually have to work alone, for example coordinators of Canadian Community Adult Learning Councils (Gray, 2004), or teachers (Barab, MaKinster, & Scheckler, 2003).

Several researchers have found evidence for the importance of legitimacy and issues of power for LPP (e.g. Davies, 2005). The case study by Back (2011) on the LPP process of two newcomers in an Andean folkloric music band also supports the argument that legitimacy is the key to opportunities for learning. In this study, the cases of two band members were explored using ethnographic methods. Both newcomers had started with similar conditions regarding their experience as musicians and also their lack of speaking skills in the traditional Quichua language in which the band's music was performed. This language is closely related to the traditions of the long suppressed Otavalo culture, and it made a significant difference for the integration process of both members, to what extent they felt connected to this culture. Although both persons were willing to learn the language and to engage in the music business of the group, only the person who showed a personal connection to the Otavalo culture was fully integrated and granted access to opportunities for learning what was necessary for the CoP, namely the language, while the other person was marginalized.

Only little research has focused on the second part of the participative learning approach, the peripheral participation. An example of such a study is provided by Lambson (2010), who investigated the participative learning process of three newcomers in a CoP of literacy teachers who met regularly to discuss ideas about how to improve their teaching and to share classroom stories about teaching experiences. Legitimation was in this CoP never a problem for the three newcomers who were new teachers at the school. In the beginning, the two complete newcomers participated mostly passively and described feelings of insecurity because of their newcomer status, while the one who had already known the other teachers for some time, participated more actively. Over time, all of them showed more active forms of participation and reported that they felt more secure in their role as a teacher. Also changes in newcomers' talks was observed from mere reports of how they had designed the teaching in a class in the beginning towards a focus on students' reactions to the ways of teaching they tried.

In this CoP, also certain ways of facilitation of the LPP process were observed. The CoP had a leader who took care that every member was heard and often asked newcomers directly when she felt that they could contribute something. She engaged all members, including the newcomers, in conversations with herself and the rest of the community, shared her own experiences on questions relevant to newcomers, and provided access to relevant information and resources. This study points to the importance of the question of what exactly happens during peripheral participation. However, this question has attracted little attention so far and lacks systematic analysis.

Both studies are representative examples for current research on participative learning that focuses mostly on the epistemic parts of participative learning (newcomers' engagement in CoP practices) and only marginally explores the social aspects of the integration into a CoP.

Social Network Perspective on Learning as Participation

In the context of the LPP approach, process and outcome of learning are closely related, interdependent, and are necessary conditions for each other. There is no end of the learning process which would facilitate the measurement of a learning outcome. This is especially the case, because not only the individual person learns, but also the CoP as a whole develops during the learning processes of its individual members. For the learner in the CoP, learning is not only the acquisition of knowledge, but concerns identity. These conditions do hardly allow for applying "standard quantitative methodology" and we see that most research on CoPs and learning in CoPs is ethnographic and of qualitative nature so far and only few quantitative studies have been conducted (Nistor & Fischer, 2012). Another reason for the sparseness of quantitative methods on the topic might be that researchers have mostly focused on the epistemic/practice part of CoPs in their studies and practices are highly complex, often seemingly unique and require in-depths analyses. However, the social aspects of learning in CoPs is also very important within this concept and this aspects better allows for the use of quantitative measures.

A corresponding statement that relations between newcomers and old-timers in the CoP and their changes over time are central aspects of LPP, can also be found in the original book on LPP (Lave & Wenger, 1991, p. 59/60):

> "[...] learning as increasing participation in communities of practice concerns the whole person acting in the world. Conceiving of learning in terms of participation focuses attention on ways in which it is an evolving, continuously renewed set of relations; this is, of course, consistent with a relational view, of persons, their actions, and the world, typical of a theory of social practice."

On this background we suggest that SNA is suitable way for the investigation of LPP in CoPs, because it has the potential to extend our current knowledge on LPP in ways other methodological approaches cannot provide.

Social Network Theory as a Methodology to Analyze LPP in CoP

The basis of all social networks are relations between individuals, called ties. Individuals are usually persons, but can also be organizations or larger groups of persons. Two individuals in a social network can be linked by three different types of ties (Borgatti, Mehra, Brass, & Labianca, 2009). (1) a tie can be based on similarities; this can be being member of the same group or sharing an attribute. (2) a tie can be based on social roles within the social network; these roles can e.g. be based on forms of cohabitation like being friends or colleagues, but also based on affective (e.g. liking, or hating each other) or cognitive states (e.g. knowing something about someone). (3) a tie can be based on interactions between individuals, e.g. if two individuals talk to each other.

Additionally, ties can either be seen as absent, weak or strong (Granovetter, 1973). The strength of a tie between two individuals depends on the amount of time they spend with each other, the intensity of their emotional attachment, or the reciprocal services. Granovetter (1973) argues, and his hypotheses have been supported by empirical findings (e.g. Granovetter, 1983) that individuals live in close knit social networks with strong ties (like family and friends), which support them in many ways. However, individuals have also several weak ties (acquaintances), which are bridges to the otherwise unrelated close knit networks of the acquaintances. These weak ties can therefore have positive attributes that cannot be found in strong ties, a prominent example is the access to new information like job offers. Wenger (1998) states that CoPs are networks of strong ties (with a focus on the practice and not only on information flow). However, it can be assumed that many forms of strong ties have developed from weak ties.

The second element of social network theory (Borgatti, Mehra, Brass, & Labianca, 2009) is the importance of structure. Social network research revealed several underlying structures and rules which affect the development of human interaction. For a social network as a whole, social network theory provides evidence that not only the composition of the network regarding important attributes of its individual actors can explain its functioning, but that the relations between the individuals are of high importance; for instance team performance is not only subject to the skills and knowledge of the team members but to a high extent of the relations between the members and the way they interact with each other. For an individual, on the other hand, her position and embeddedness in a specific social network as well as the structure of the person's own social network are also of high importance and of a great predictive value.

Social network research is, furthermore, characterized by specific types of research questions that take ties between individuals and/or the structure of a social network into account, which distinguishes it from other types of research in the social sciences, and it focuses on specific varieties of networks in different contexts, which distinguishes SNA from network research in science. A typical set of research questions deals with the consequences of social network structures, trying to predict various kinds of outcomes leading them back to variances in the social network structures of groups. Another research focus is on the antecedents of network formation, mostly focusing on dyads. This research investigates for examples factors that can predict a dyad's likelihood to form a friendships tie (Borgatti, Mehra, Brass, & Labianca, 2009).

The usually underlying theoretical mechanisms of SNA are adaption, binding, and exclusion (Borgatti, Mehra, Brass, & Labianca, 2009). It is assumed that something relevant, such as knowledge, flows between the ties of a social network. This flow can either lead to adaption in form of convergence between the connected individuals, or to a binding which can be understood similar to a chemical binding in which two elements form a new entity when they connect. Exclusion as the third mechanism applies to situations in which the number of possible relations is restricted and the formation of one tie means that other ties cannot be formed.

Methodological Foundations of Social Networks

A social network consists of two elements: The individuals, which are in SNA termed as actors or nodes, and the relations that connect them, which are usually referred to as ties or edges (Scott, 2000). Analysis and presentation of social networks relies on the one hand on sociograms which visualize the social network using dots of various shapes to display the actors and lines or arrows as representations for ties. A social network is a

multidimensional object, with each additional actor requiring a further dimension, which makes visualization in two dimensions (or three at a maximum) a challenging task. Additionally, more complex mathematical approaches are used to satisfy the requirements of the multidimensional complexity of social networks. Matrix algebra and graph-theory form the basis for SNA (Borgatti, Mehra, Brass, & Labianca, 2009).

From a researcher's perspective, two forms of SNA are possible (Hatala, 2006). An *ego*-network is the social network of one particular actor. The main actor, called ego, is the center of the network and all actors to which ego has ties appear in that network. A *complete* network analysis, by contrast, focuses on the whole network consisting of a defined set of actors and the ties between all the actors are collected. A complete social network can be analyzed on three different levels (Hatala, 2006). (1) The level of the whole network allows describing its structure. Some of the most popular measures on that level are density or cohesion, which indicate if a network is dense or loose regarding the amount of ties between the actors. Another prominent measure is centralization, which describes the network's structure as more centralized with a clear core or as rather distributed. Also modularity is a relevant measure; it indicates if a network consists of many small clusters or if all members are evenly connected to each other.

(2) On the actor level, the situation of individual actors within the network can be analyzed. On that level, the most studied concept is centrality which measures the structural importance and prominence of an actor within a social network (Borgatti, Mehra, Brass, & Labianca, 2009). Degree centrality is the simplest centrality measures and is, in short, the sum of all ties of an actor. In larger networks, it can be relevant to distinguish between global degree centrality which is the sum of all ties an actor has in the whole network, and local degree centrality which indicates the actor's centrality between several clusters of the network (Hatala, 2006). Additionally, the measurement of the tie influences the computation of centrality. There are cases in which a tie is undirected, like actor A and actor B are in the same class, and so a simple measure of centrality can be computed, but the tie can also be directed which means it can be seen differently by actor A and B, like A gives information to B, but B does not give information to A. In this case, it is necessary to distinguish between indegree-centrality and outdegree-centrality. Indegree-centrality is based on the amount of ties that are directed towards an actor, whereas outdegree-centrality is based on the amount of ties that are direct from the actor to the others. Also others, more complex measures for centrality were developed like betweenness centrality which is not only based on an actor's direct ties, but also the indirect ties that an actors has over the actors to whom she is connected (Scott, 2000). Closeness-centrality, as a further measure, is even more complex. It measures an actor's closeness to all other actors in the network (Freeman, 1979). However, degree centrality is the most common centrality measure and it allows for a more straight forward interpretation than the other centrality measures.

(3) The third level for SNA is the tie-level which usually deals with dyads and how two actors are related. For each dyad can be said, if a certain type of relation is present or not. Sometimes, it is also useful to have not yes or no outcome for a relation, but a valued relation which represents the intensity of the relation. For example, it can be asked if two persons had talked to each other, and also how long or how often they had talked to each other.

With the emergence of online social networks sites, the awareness of people about their own social network increases. Also research on social network can increase participants' awareness on their social network. Borgatti and colleagues (2009) see this as similar to the Heisenberg uncertainty principle of which researchers should be aware. However, this effect provides also potential for constructive use, when network visualizations are used to enhance actors' awareness of the structure or risky positions of individual actors in the social network, and provides possibilities to intervene. The work of Yoon (2011) builds on that idea.

As we argued before, participative learning in the form of LPP can be understood as an increase of newcomers' participation within a social network. Therefore, the social network as a whole is not as relevant, as the person-level and the tie-level are for the investigation of learning. On the person-level, one of the most used indicators is centrality, a measure, which is also relevant for studying participative learning. Centrality measures to what extent an individual is related to other persons in the network. So, successful LPP can be operationalized as an increase of newcomer's centrality. But also the tie-level can be relevant, to study which factors have influence of a tie's probability to turn into a certain type of relation.

Examples of Applied Social Network Analysis to Study Learning

In the Learning Sciences SNA is not yet a standard approach, but also not unknown. Some studies can be found that are based on this methodology. CSCL, for example, has been studied in formal settings using SNA approaches. Nurmela, Lethinen, and Palonen (1999), for instance, applied a social network approach in a study on students who participated in a university course on educational psychology and worked on a learning platform. The students worked in dyads for a semester and had to complete weekly assignments. The online learning environment allowed the joint writing of assignment texts and commenting to texts of other students' assignment texts. A tutor encouraged interaction across the student dyads. SNA was used to analyze logfiles of the joint writing process and the interaction of students via the commenting tool. The social network approach

provided a way to display the interaction of students in a structured way, to identify central actors in the commenting process, to show inequalities between the dyads and to identify the key documents of joint engagement in the commenting process.

Another social network study on CSCL by Reffai and Chanier (2000) investigated cohesion in a large scale distance-learning course on French as a foreign language. Cohesion, in the sense of regular participation and commitment to the class, is considered as highly important for collaborative learning as it is the basis for the learners' motivation and ability to benefit from the teaching approach. Similar to the study by Nurmela and colleagues (1999), social network measures were applied to make the complex communication in the virtual learning platform visible, finding cliques as well as central or isolated students in the discussions. Reffai and Chanier (2000) conclude that such measures can be very useful for tutors and lecturers to monitor the complex communication and a monitoring tool based on SNA would be a helpful tool.

In contrast to formal learning settings, SNA has rarely been applied in informal settings to study learning. As we argued, especially for these settings and when studying learning as participation, SNA has high potential to push research forward. To illustrate this kind of research, we will describe two studies we conducted with the aim to bring further our understanding of learning as participation by focusing on the social aspects of this approach.

Using SNA to Study the Role of Integration Tactics on Newcomers Participation

As our review of research on legitimate peripheral participation has shown, the role of old-timers facilitation of the LPP process has rarely been explored although scattered findings indicate its importance. Therefore, we conducted a study with the goal to investigate how old-timers of CoPs structure the participative learning process of newcomers (Eberle, Stegmann, & Fischer, 2014).

Our sample consisted of the members of 14 faculty student councils, which are CoPs of university students at one particular faculty sharing the goal to improve student life and education from the student perspective. We implemented a mixed-method approach including SNA to tackle the question. To capture how old-timers' facilitate and influence the participative learning process of newcomers in their CoP, we interviewed an experienced member of each CoP and extracted several distinct integration strategies using qualitative content analysis. We then assigned a quantified value for the use of each of the integration strategies within the individual CoPs to allow for a regression analysis to explore the relation of these integration tactics to the state of participative learning in the newcomers of the CoPs.

SNA methodology was applied to identify newcomers' states of participative learning. We asked all members to fill a social network questionnaire and to indicate on a Likert-scale for each of the other members to what extent they had collaborated with them during the last three months. We received a collaboration network with valued ties for each of the CoPs and decided to take the indegree-centrality measure of the newcomers as indicator for their state of participative learning, taking only the ratings of old-timers into account. We opted for a degree-centrality measure as its interpretation is the easiest compared to other centrality measures. The decision for using the indegree-centrality was based on the finding that old-timers rated much more conservative then newcomers who probably lacked the experience to estimate the intensity collaboration can have in the particular CoP as it is perceived by experienced members.

HLM models based on our measures for integration tactics of the CoPs and individual states of participative learning of their newcomers showed that certain tactics (especially the tactic to make specific community knowledge accessible to the newcomers) seem to be influential during the LPP process of newcomers across several CoPs. Our focus on the social aspects of LPP using SNA gave as the possibility to investigate LPP across the very specific individual conditions in CoPs which would have hardly been possible if we had focused on epistemic aspects like identifying newcomers' use of certain CoP specific practices. In this study, SNA was the key to go beyond case studies within one CoP and to make comparison between participative learning processes in different CoPs possible.

Using SNA to Study the Role of Integration Tactics on Newcomers Participation

In our second study (Eberle et al., 2013), we wanted to explore if the most promising integration tactic we found in faculty student councils is also influential in another type of CoP and investigate the detailed effects of this tactic more in depth. The integration tactic we looked at was the provision of access to community knowledge which can be implemented in several ways from written information to oral instructions, and from information about community practices to information about individual community members.

We set up a quasi-experimental study in a scientific community and studied the differences that occurred between members who had received supportive access to community knowledge and those who did not receive this support measure during a scientific event. Our support measure was a brochure with information about the other members.

This study was solely based on an SNA approach and several ways of data collection about the social network structure and ties between individuals were implemented: (1) we tracked the interaction between the

CoP members during the scientific event using RFID devices that captured who interacted with whom, (2) we distributed SNA questionnaires asking the members with whom they had collaborated before the event and with whom they planned to collaborate afterwards, and (3) we run a google scholar search 1.5 years after the event to identify coauthored publications. The analysis focused on the emerging ties between the members and revealed that our intervention had an influence on the interactions between members at the event with those in the group who received access to community knowledge interacting with less other members than those without access to community knowledge, leaving more time for longer interactions; longer interaction times, furthermore, lead to a higher probability of identifying potential for future collaboration.

Our SNA based study design allowed for a very detailed data collection of occurring collaboration and interaction that would not have been possible with traditional approaches like video studies. We were able to follow the participants' behavior during a 4 day period during the whole day and in a wide locational range without disturbing or hindering spontaneously occurring interaction. Through the combination of the behavioral measures for interaction and the SNA questionnaire, we were able to capture natural interaction behavior as well as the individuals' interpretations of these interactions had for the observed persons. We were even able to identify to what extent intended behavior had culminated in productivity and the production of artifacts that are of high value for the particular community. The triangulation of these different aspects of embeddedness in the social structures of the CoP allowed us to identify the current state of a person within the CoP in a sophisticated multi-facetted way.

Conclusions

We have argued that research based on the metaphor of learning as participation, which is closely related to the approach of LPP in CoPs, has so far focused mostly on the development an re-instantiations of practices – which can be seen as the epistemic aspect of participative learning. We suggest taking the social aspects, namely the integration of newcomers into the social network of a CoP more into account. From a methodological perspective, the two studies we presented on learning as participation using an SNA approach are rare cases of quantitative research in the field of learning in CoPs (Nistor & Fischer, 2012). Case studies are so far the dominant research method for participative learning in CoPs. An important reason for that probably is that each CoP is a complex and distinct constellation of factors; some practices are only relevant in that specific CoP and also the way how members mutually engage do differ remarkably between CoP. Therefore, it is a critical question to what extent findings from one CoP can be transferred to another or if these results are only valid for the specific context. However, learning has always be seen as a complex process with important individual and contextual influences, no matter under which paradigm it is investigated. It is a main characteristic of quantitative research that it aims at identifying re-occuring and hence maybe generalizable patterns that are valid across individual and contextual characteristics. Such general patterns are still missing in research on participative learning in CoPs. In contrast to this, the use of SNA that allows for research across individual CoPs can also show which aspects are indeed specific for certain CoPs and cannot be transferred. At the moment, we just have a scattered body of research that neither allows for conclusions about CoPs and participative learning in general, nor for the argumentation that findings are specific to certain CoPs or types of CoPs. Instead, the accumulated body of research is very diverse and full of different understandings of the core concept. After more than 30 years of research in this field, it seems not too early to explore the concepts also more systematically including also quantitative means.

Taking a look at the social aspects of participative learning instead of only on the epistemic aspects seems to be a promising approach in this direction. Means of SNA have turned out to be a promising approach for this endeavor. The focus on social aspects identified by applying means of SNA allow for comparison across specific epistemic settings and individual cases. It allows research in so far under-investigated areas of learning, especially in informal learning settings that can hardly be approached using traditional methodology because it is impossible to predict when and where informal learning opportunities occur and what they will be about. Developing technological means like the use of RFID devices attached to individuals in physically co-present communities allow new approaches to collect data in these areas that have mostly been "un-researchable" until now. Also developments in data analysis methods, like dynamic SNA and other means of capturing network evolution, allow for new research questions and extension of the current focus. The combination of newly developing and more traditional methods like questionnaires and interviews seems especially promising, as it not only allows for capturing different types of data but also for triangulating and validating these data.

However, only very few of the possibilities that SNA theory and methods offers have been applied in the Learning Sciences so far. Not only research on learning on CoPs could benefit from SNA, disciplines like education and the learning sciences have not yet identified the full potential of this approach for the understanding learning and social factors related to education. However, the most beneficial step would be to integrate both perspectives, the epistemic and the social perspective for further investigation of learning under the participation metaphor, by looking at emerging and changing practices in communities in relation to the emerging and dynamically changing participation processes of individuals and groups of individuals.

References

Eberle, J., Stegmann, K., & Fischer, F. (2014). Legitimate Peripheral Participation in Communities of Practice—Participation Support Structures for Newcomers in Faculty Student Councils. *Journal of the Learning Sciences,* doi:10.1080/10508406.2014.883978.

Eberle, J., Stegmann, K., Lund, K., Barrat, A., Sailer, M., & Fischer, F. (2013). The development of collaborations in a scientific community – evidence from an experiment using sociometric RFID devices. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 1 – Full Papers & Symposia* (pp. 169-175). International Society of the Learning Sciences: Madison, WI, USA.

Back, M. (2011). Legitimate Peripheral Participation and Language Learning: Two Quichua Learners in a Transnational Community. *Language Learning, 61*(4), 1039-1057.

Barab, S.A., & Duffy, T.M. (2000). From Practice Fields to Communities of Practice. In D. H. Jonassen (Ed.), *Theoretical foundations of learning environments*. Mahwah, NJ: Erlbaum.

Barab, S.A., MaKinster, J.G., & Scheckler, R. (2003). Designing system dualities: characterizing a web-supported professional development community. *The Information Society, 19*(3), 237-256.

Borgatti, S.P., Mehra, A., Brass, D.J., & Labianca, G. (2009). *Network Analysis in the Social Sciences. Sciences, 323*, 892-895.

Boyd, P., & Lawley, L. (2009). Becoming a Lecturer in Nurse Education: the workplace learning of Clinical Experts as Newcomers. *Learning in Health and Social Care, 8*(4), 292–300.

Campbell, M., Verenikina, I., & Herrington, A. (2009). Intersection of trajectories: a newcomer in a community of practice. *Journal of Workplace Learning, 21*(8), 647-657.

Collins, A. (2006). Cognitive Apprenticeship. In R.K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 47-60). Cambridge: Cambridge University Press.

Davies, B. (2005). Communities of Practice: Legitimacy not Choice. *Journal of Sociolinguistics, 9*(4), 557-581.

Fuller, A., Hodkinson, H., Hodkinson, P., & Unwin, L. (2005). Learning as Peripheral Participation in Communities of Practice: a Reassessment of Key Concepts in Workplace Learning. *British Educational Research Journal, 31*(1), 49-68.

Granovetter, M (1973). The strength of weak ties. *American Journal of Sociology, 78*, 1360-1380.

Granovetter, M (1983). The strength of weak ties: a network theory revisited. *Sociological Theory, 1*, 203-233.

Gray, B. (2004). Informal learning in an online community of practice. *Journal of Distance Education, 19*(1), 20-35.

Hatala, J.-P. (2006). Social network analysis in human resource development: a new methodology. *Human Resource Development Review, 5*, 45-71.

Hoadley, C.M. (2005). The shape of the elephant: scope and membership of the CSCL community. In T. Koschman, D.D. Suthers, & T.-W. Chan (eds.), Computer supported collaborative learning: the next 10 years! Proceedings of the international conference on computer supported collaborative learning 2005 (pp. 205-2010). Mahwah, NJ: Lawrence Erlbaum Associates.

Lambson, D. (2010). Novice teachers learning through participation in a teacher study group. *Teaching and Teacher Education, 26*(8), 1660–1668.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. Cambridge: Cambridge Univ. Pr.

Nistor, N. & Fischer, F. (2012). Communities of practice in academia: testing a quantitative model. *Learning, Culture and Social interaction, 1*, 114-128.

Nurmela, K.A., Lehtinen, E., Palonen, T. (1999). Evaluating CSCL log files by Social Network Analysis. In C. M. Hoadley and J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference*. Palo Alto, CA: Stanford University, 434-444.

Reffai, C., & Chanier, T. (2000). *How social network analysis can help to measure cohesion in collaborative distance-learning*. Dordrecht, The Netherlands: Kluwer.

Sfard, A. (1998). On two metaphors for learning and the danger of choosing just one. *Educational Researcher, 29*(4), 4-13.

Wenger, E. (1998). *Communities of Practice. Learning, Meaning, and Identity*. Cambridge: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston, Mass.: Harvard Business School Press.

Yoon, S. (2011). Using social network graphs as visualization tools to influence peer selection decision-making strategies to access information about complex socioscientific issues. *Journal of the learning sciences, 20*(4), 549-588.

Designing Critique to Improve Conceptual Understanding

Elissa Sato and Marcia C. Linn, University of Berkeley, California, 4523 Tolman Hall, Berkeley CA 94720
elissa.sato@berkeley.edu, mclinn@berkeley.edu

Abstract: Students become entangled in their varied scientific ideas and struggle to reconcile their understanding with ideas encountered in instruction. This design-based study with a sixth-grade technology-enhanced inquiry science unit on global climate change investigates how critique can support students in refining their conceptual understanding. Specifically, the study investigates whether students' ability to benefit from critique is impacted by the complexity of the critique artifact. Findings show that students can equally benefit from critiquing explanations of varying complexity when guided to consider a range of alternative ideas during critique The results show the value of designing critique to support students in distinguishing among their own and alternative ideas. Case studies illustrate how students engaged with opportunities provided by the guidance, and indicate areas where further research is necessary to refine the design of critique as a means to support conceptual learning in science.

Rationale

Students become entangled in multiple, often conflicting ideas about scientific phenomena as they interact with the natural world and struggle to distinguish new ideas from existing beliefs (e.g., Clark, 2006). Both children and adults resist and discount evidence that contradicts their existing beliefs (Chinn & Brewer, 1993). Yet, citizens need to develop the ability to use scientific evidence to critique ideas of others and to interpret critiques of their own ideas. Efforts to date offer some promise for critique but also reveal the need for clarification of how, when, and why critiques are beneficial for conceptual learning (e.g., Shen, 2010). This study seeks to advance our understanding of critique by comparing two approaches to designing critique.

In designing critique, we draw on the constructivist knowledge integration (KI) framework (Linn & Eylon, 2006) that addresses the difficulties students have in making sense of their multiple, conflicting ideas. KI calls for building on the repertoire of ideas students develop in their lives by designing inquiry experiences that support students in considering alternatives and refining their conceptual repertoire. However, students' ability to distinguish among alternatives during critique may be dependent on the complexity of the critique artifact. The first approach draws on the notion of the zone of proximal development (Vygotsky, 1978), which suggests that students are most likely to benefit when the learning task is designed to align with their prior knowledge such that the task is accessible and allows students to make progress with appropriate guidance. Guiding students in critiquing a normative response that is incomplete yet slightly more sophisticated than their current explanation could support students in distinguishing among ideas without being overwhelmed by complexity.

However, alternative perspectives such as desirable difficulties in psychology (Bjork, 1994) and productive failure in mathematical problem solving (Kapur & Bielaczyc, 2012) suggest that reducing the complexity of cognitive tasks may have a detrimental impact on student learning by deemphasizing the need to distinguish among ideas in their conceptual repertoire. From the KI perspective, conceptual critique involves distinguishing among normative and non-normative ideas. Critiquing a slightly more sophisticated normative response may not support students in this process because the non-normative ideas are not explicit. Students may be content with addressing the more obvious flaws and neglect to reflect on the range of ideas. Thus, critiquing a complex response with a mix of normative and non-normative ideas may be more successful in supporting deep understanding by prompting students to reflect more holistically on their conceptual repertoire. Our study therefore seeks to address the following research questions:

1. How do students benefit overall when they critique (a) an incomplete explanation with normative ideas to identify a missing idea (*incomplete*) or (b) an incomplete explanation combining normative and non-normative ideas to identify a non-normative idea (*non-normative*)?
2. How do students' ideas, as expressed in their explanations, shift in response to critique?

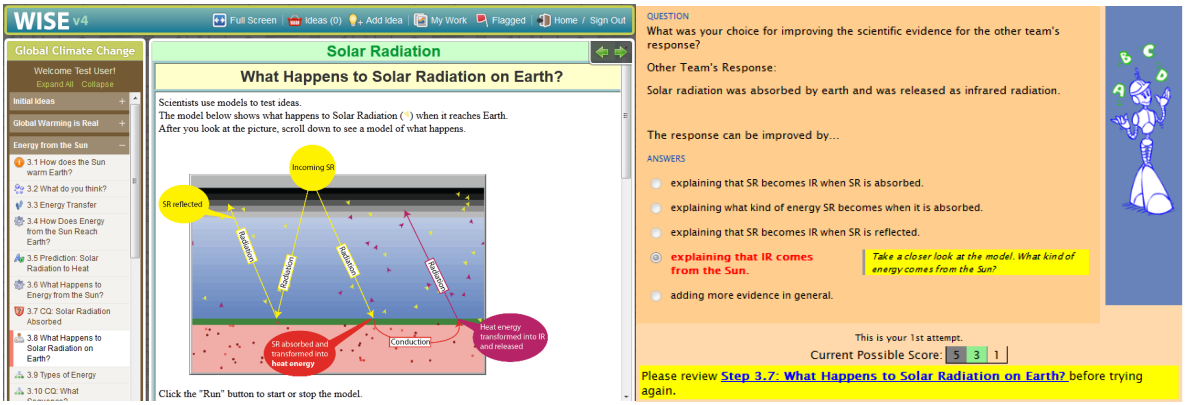
We hypothesized that we would observe significant differences between conditions in students' learning gains if the potential benefit of critique depended upon carefully designing an accessible or *desirably difficult* critique artifact that was aligned to the students' prior knowledge. On the other hand, we hypothesized that students in both conditions would make comparable progress in their learning if the potential benefit of critique were less dependent on the complexity of the critique artifact and more dependent on whether students were appropriately supported in considering alternative ideas.

 © ISLS

Methods

A sixth-grade technology-enhanced earth science curriculum unit, Global Climate Change (GCC, Figure 1a), was developed in the Web-based Inquiry Science Environment (WISE, Linn, Davis, & Bell, 2004) using the KI perspective. In GCC, students grapple with the complex energy mechanisms driving changes in global climate through a series of interactive NetLogo simulations (Svihla & Linn, 2012). Students are provided with multiple opportunities to explain causal subsets of this complex system before generating an integrated explanation of the overall phenomenon. They investigate how factors such as greenhouse gases impact energy transformation and how that in turn impacts global temperature trends. Student explanations were coded for the sophistication of their mechanisms. We focused on an explanation targeting an energy transformation process critical to understanding the phenomenon of global climate change. The GCC unit was completed by 68 middle school students working in pairs taught by the same teacher, who had taught previous versions of the unit.

Activity sequence. Student pairs were randomly assigned to one of two promising approaches for aligning critique artifacts to students' ideas (Fig, 2). Based on the design research paradigm, the study sought to investigate whether one condition is better than the other. In the activity sequence students generated an initial explanation, critiqued and revised an assigned explanation, received conceptual guidance on their critique, then critiqued and revised their initial explanation (Figure 2). The design focused on encouraging students to revisit evidence steps (e.g., simulations), which has been correlated with learning gains (Svihla & Linn, 2012), and to discuss and negotiate alternative ideas presented through the critique artifact and critique choices.



(a) (b)

Figure 1. (a) The WISE *Global Climate Change* unit. (b) The guidance checkpoint step.

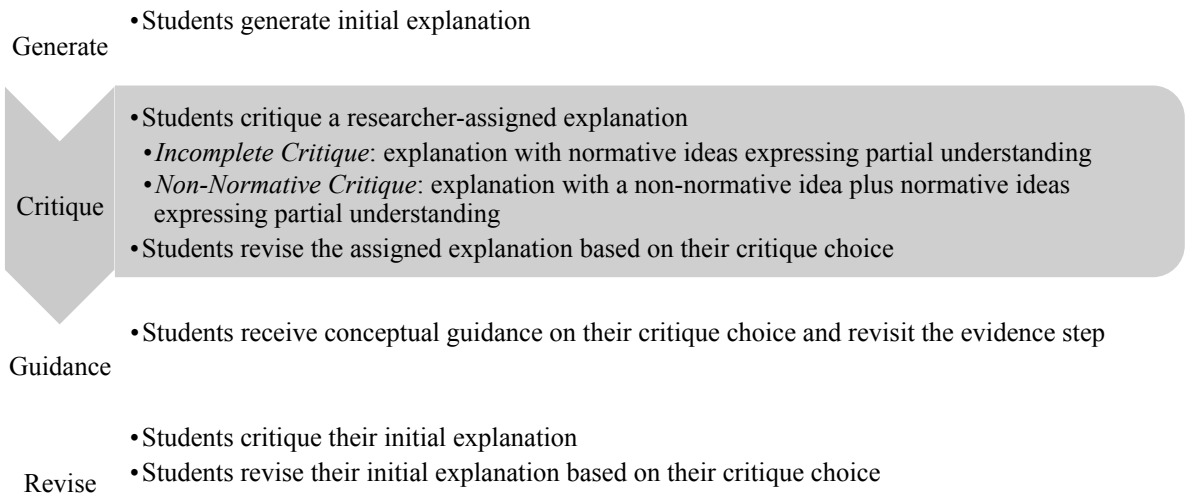


Figure 2. Outline of the overall activity sequence.
The shaded step indicates where the curriculum design differed between the two versions.

Critique artifacts. For each condition, three critique artifacts were designed by the researcher based on the analysis of responses collected during previous classroom implementations. The researcher assigned an explanation that expressed partial understanding and was slightly more sophisticated than the initial explanation generated by the students. The *incomplete* group critiqued an explanation containing only normative ideas. The *non-normative* group critiqued a modified version of the *incomplete* explanation with a non-normative idea. During critique, students in both groups selected a science content critique from among several alternatives. This entailed distinguishing among alternatives. They then revised the critiqued explanation based on their choice (Table 1).

Table 1: Critique step guidance.

Feature	Prompt
1) Critique Artifact	The other team's response: Solar radiation was absorbed by earth and released as infrared radiation.
2) Critique of Surface Features	Score this response for spelling, grammar, and punctuation: • Very good: No spelling, grammar, or punctuation errors • Good: Few spelling, grammar, or punctuation errors • Not So Good: Many spelling, grammar, or punctuation errors
3) Critique of Science Content	What needs to be changed in the response to improve the scientific evidence? The response can be improved by… • Explaining what kind of energy SR becomes when it is absorbed. • Explaining that IR comes from the Sun. • Explaining that SR becomes IR when SR is reflected. • Explaining that SR becomes IR when SR is absorbed. • Adding more evidence in general.
4) Revision of Critique Artifact	Change and improve the other team's response based on your choice above.

Note. The design of the critique guidance was adapted from a previous study (Sato & Linn, 2011). The guidance was consistent across conditions. The conditions differed in the explanations they were assigned to critique and in the science content critique choices displayed. The same guidance was provided in the revision step.

The critiques for scientific evidence were specific to the explanation type and targeted missing ideas and connections among ideas in the *incomplete* group, or non-normative ideas and connections in the *non-normative* group. Critique choices were calibrated such that both conditions considered the same range of alternatives for a given explanation.

Guidance checkpoint. Both conditions had additional opportunities to consider the same alternatives when they received automated conceptual feedback on their critique choice at a guidance checkpoint. They received a guiding question and were prompted to revisit a critical step to reevaluate the evidence (Figure 1b). During the guidance checkpoint, students were discouraged from mindless guessing with choices that changed order between attempts and a diminishing score structure. During revision, students were also prompted to draw on their critique experience by applying the same criteria to their explanation. They were also prevented from referencing the critique step so that they would not copy ideas from the critique artifact.

Data

Student work formed the core data source; the unit of analysis was the dyad. Student responses were coded using a rubric based on the KI framework, which rewards coherence of ideas as represented by the number and complexity of connections students make between their ideas (see Table 2). Ten pairs whose initial explanations already demonstrated complex understanding were removed from the analysis.

Table 2: Knowledge Integration rubric used to score students' original and revised explanations.

Explanation Prompt: Where did infrared radiation (IR) come from in the model? Give as much detail as you can.		
Score	**Description**	**Student Examples**
1 (Irrelevant)	No answer or irrelevant answers	I don't know
2 (No Link)	Non-normative ideas or links	It came from the sun and space.
3 (Partial Link)	One relevant and normative idea	IR came from conduction, under the earths crust. The Solar Radiation transforms into heat energy that bounces off earths crust and is trapped by the greenhouse gases and unable to escape earths atmosphere.

4 (Full Link)	Scientifically valid and fully elaborated link between two relevant and normative ideas	It comes from heat energy when heat energy is released it goes into the Infrared radiation, so it becomes heat energy.
5 (Complex Link)	At least two links among three or more relevant and normative ideas	Some solar radiation is reflected back into space, and some is absorbed. The SR that is absorbed becomes heat energy, and heats up the Earth. It is in there for a while, and is eventually is released back into the atmosphere as infrared radiation.

Note. Examples are actual unedited responses by students.

Impact of the GCC Unit on Overall Learning Gains

Students made significant pretest to posttest gains across conditions (Table 3). There was no significant effect of condition after controlling for pretest scores ($F(1,27)=0.45$, $p>.05$). Thus all students benefitted from the unit, including the critique activities.

Table 3: Means and standard deviations for pre and posttest by condition.

		Pretest		Posttest		t	Effect Size	*p*
	N	M	SD	M	SD		*d*	
All	29 pairs	2.97	0.19	4.00	0.93	5.68	1.54	<.001
Incomplete	11 pairs	3.00	0.00	3.81	0.87	3.11	1.32	<.05
Non-Normative	18 pairs	2.94	0.24	4.11	0.96	4.75	1.67	<.01

Weighing Alternatives Effective for Supporting Revision in Both Conditions

On the embedded assessments, there was significant improvement from the students' original to revised explanation across groups (Table 4). The critique guidance helped students in both conditions revise their explanations, with medium effect sizes. There was a slight trend for the non-normative condition to make larger gains but no significant differences between conditions after controlling for pretest scores ($F(1,27)=0.05$, $p>.05$).

Table 4: Means and standard deviations for original and revised explanation scores by condition.

		Original		Revised		t	Effect Size	*p*
	N	M	SD	M	SD		*d*	
All	29 pairs	2.72	0.75	3.24	1.02	3.55	0.58	<.005
Incomplete	11 pairs	2.91	0.83	3.69	1.12	2.89	0.46	<.05
Non-Normative	18 pairs	2.61	0.70	3.17	0.99	2.56	0.66	<.05

Shifts in Students' Ideas

We analyzed students' initial and revised explanations for a shift in use of science ideas. Students' explanations were coded for scientifically valid ideas that were targeted by the explanation prompt, as well as non-normative and partially normative ideas used to assign students to specific critique artifacts (Figure 3). Ideas were coded as partially normative when their mechanistic depth was missing details that were targeted by the explanation prompt, but were not non-normative per se.

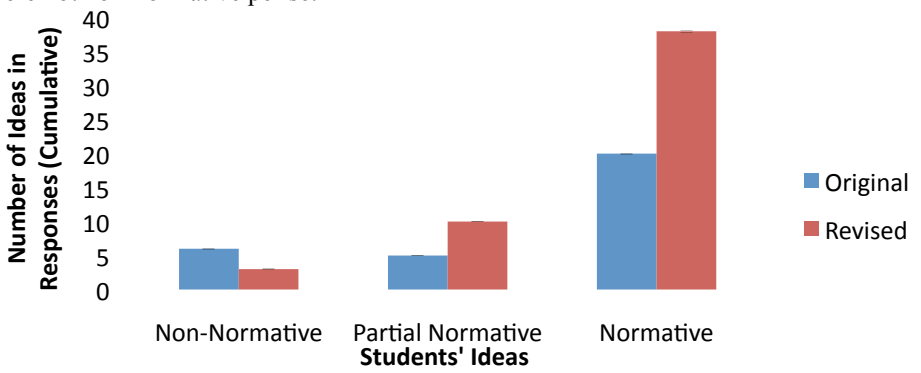


Figure 3. Shifts in students' ideas expressed in original and revised explanations across conditions.

Although not an exhaustive list, the ideas were selected for coding based on their prevalence in student responses collected during previous implementations of the unit. There was a significant gain across conditions for normative ideas, $t(29)=3.09$, $p<.01$, $d=.53$; the decrease in non-normative ideas approached significance

t(29)=-1.80, p=.083, *d*=.30; and the increase in partial-normative ideas was not significant. There were no significant differences between conditions for each category of ideas. These results provide support for our hypothesis that critique supports students' conceptual learning of scientific phenomena by guiding then to consider a range of alternatives. Results were not influenced by the complexity of the critiqued artifact.

Value of Multiple Opportunities to Reconsider Alternatives

To investigate the general impact of the activity sequence on students' success in critique and revision of explanations, we analyzed students' science content critique and revision during the critique step prior to the guidance checkpoint and during the revision step after guidance checkpoint (Figure 4). Students' critiques were coded as a success if they selected the correct science content critique. Revisions were coded as successful if they led to a gain in KI scores relative to the initial score.

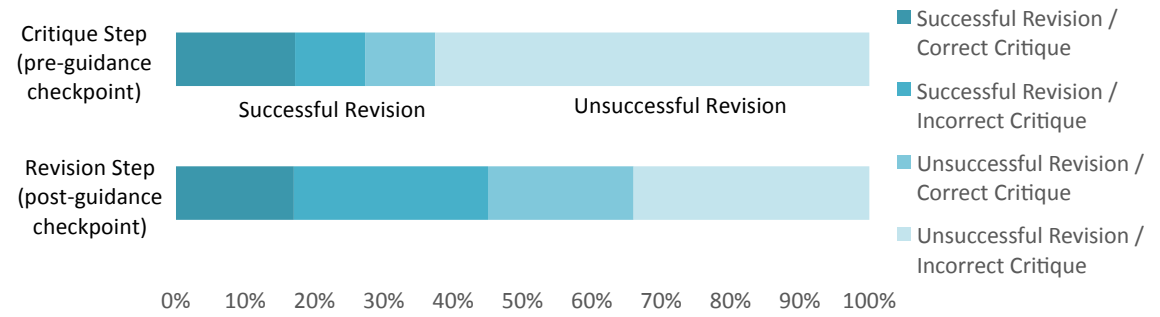


Figure 4. Frequency of Successful and Unsuccessful Revision and Critique of Assigned Explanation and Students' Explanation by Step across Conditions

Overall, the number of students who successfully revised either the assigned or their own explanation increased from the critique to the revision step. This is an encouraging finding, given that conceptual revisions are especially difficult for students, even if they receive direct feedback on the written artifact to be revised (Cho & MacArthur, 2010). In this study, students only received conceptual guidance on their critique choice. During the critique step, only 27% of all students made a successful revision of the sample explanation based on their choice. Critique was challenging for students such that 72% of them selected an incorrect critique. However, 14% of students who selected an incorrect critique were still able to improve the critiqued explanation. Grappling with critique which involves considering alternative ideas, even when unsuccessful, may still support students in making productive revisions. Following the guidance checkpoint, more students (45%) made a successful revision of their own explanation. Although 62% of students still struggled with critique, a greater percentage of those students (44%) made a successful revision of their own explanation. These positive shifts indicate that, at least for some, the guidance checkpoint was a valuable opportunity to reconsider their ideas in light of alternatives and make further progress after the critique step.

Preliminary 2x2 Chi square tests suggest that there may be a significant association between critiquing and revising the target artifact (the assigned explanation) during the critique step ($\chi 2(1)=6.74$, $p<.05$). Essentially, students were 10.35 times more likely to make a successful revision of the critique artifact if they selected the correct critique during the critique step. However, there was no significant association between successful critique and successful revision of the students' own explanation during the revision step ($\chi 2(1)=.00$, $p>.05$). The decoupling of critique and revision following the guidance checkpoint supports the idea that receiving conceptual guidance and an opportunity to revisit a key visualization allowed students to make successful revisions of their own explanation despite their struggles with critique.

Classroom observations. To further examine how students engaged with the various steps comprising the activity sequence, we used classroom observations and video records. In this paper, we characterize student engagement as the types of discussions students had with each other and their interactions with the activity scaffolds. During critique, we observed that some students seemingly guessed when initially selecting their science content critique and did not discuss alternatives until prompted to revise the critiqued explanation, while others discussed the critique choices during selection. During the guidance checkpoint, some were frustrated by the complexity of selecting among plausible alternatives and engaged in guessing behavior, whereas others leveraged the additional opportunity provided by feedback to reassess their understanding or to request help from the instructor or researcher.

To illustrate the kinds of engagement observed in the overall data corpus, we present descriptions and transcribed excerpts of video records. The video data suggest that the design can provoke opportunities for students who may not otherwise engage in negotiation and reconsideration of ideas, as well as for students who are already doing so. However, how to ensure that such opportunities are leveraged by students remains an open question, as we discuss below. In this paper, we focus on the guidance checkpoint, because it was intended to

serve as a pivotal opportunity for students to reconsider their own ideas and their assessment of alternative ideas during critique.

Capitalizing on Opportunities: Collaborative Sensemaking and Reflection on Ideas

Janelle and Ida took turns controlling the computer and answering prompts. They were jointly engaged with the unit, discussing science content and co-constructing responses to prompts. They also asked each other for confirmation while commenting on ideas with questions such as "We're OK, right?" and "How's that?" before finalizing their work. Their engagement pattern persisted throughout the activity sequence, with both partners commenting on the critique choices. Their aptitude for collaborative sensemaking and deliberation raises the question of whether the activity design adds value to their learning process. The transcript below suggests that their existing orientation allowed them to capitalize on opportunities afforded by the activity design and further refine their ideas. Prior to this moment in the guidance checkpoint, they had worked their way through critique, assessing each critique choice with regard to its scientific validity, but without justifying why by referencing relevant ideas (e.g., "That's not true."). They made a successful critique and revision of the critique artifact (Table 5), but during the guidance checkpoint, Janelle argues for a different critique choice ("needs more evidence in general").

1 Janelle: I think "adding more evidence in general" because they didn't really explain where the energy comes from or what it transforms into. (J chooses the choice and submits; it's wrong)
2 J: Oh.
3 Ida: Wait.
4 J: Sorry.
5 I: Oh wait we have to review it. (I goes back to simulation step and reviews text preceding the simulation while saying, "Blah, blah, blah.")
6 J: OK, go back to the [guidance checkpoint] step. (I continues to and starts simulation. I and J watch it silently for 11 seconds, then I goes back to guidance checkpoint. I reads through options and evaluates each with "That's not true," etc. with J watching)
7 J: (Sighs) Wait, "Explaining that IR comes from the Sun" (.) But not directly. (.) It doesn't come, like, directly though.
8 I: Yeah.

In this example, the activity design provides an opportunity for the students to further refine their understanding because the unsuccessful attempt during the guidance checkpoint prompted them to revisit and review the simulation (Line 5) and re-evaluate the choices they had previously evaluated during critique (6-7). Unlike during critique when they had evaluated the choices without justification, this time Janelle elaborated why she agreed or disagreed with the choice (7, also 1 prior to receiving guidance). Similar instances were observed elsewhere in the corpus during the activity sequence where they reassessed the content more carefully after an initial attempt to select an alternative. Although their revision of the critique artifact indicates complex understanding of the target ideas (Table 5), they chose to elaborate on albedo's role in the process, an untargeted but relevant idea, when revising their own explanation. Their revision was the only response in the data corpus that demonstrated increased sophistication but focused on the amount of energy transformation (using albedo) rather than on the kind of energy SR becomes. Their discussions and actions during the activity sequence provide evidence for how the activity's design can create opportunities for students to reconsider their ideas, which in Janelle and Ida's case enhanced their engagement with the science content and deepened their understanding.

Table 5: Critique artifacts, critiques, and revisions by Janelle and Ida

Critique Step	Artifact (Assigned Explanation)	The creation of infrared radiation began when the solar radiation comes from the Sun. Some radiation is absorbed or reflected. The ones that were absorbed goes through Earth and eventually come back out of the Earth and becomes infrared radiation. (KI Score = 4)
	Critique	The response can be improved by… explaining what kind of energy SR becomes when it is absorbed. (Correct choice for critique)
	Revision	The creation of infrared radiation begins when the solar energy radiates from the Sun. Some radiation is absorbed and/or reflected. The rays that were absorbed travel through Earth's lithosphere, transform into thermal energy, and eventually exits out of the Earth's surface as infrared radiation. (KI Score = 5)
Revision Step	Artifact (Original	When solar radiation comes in contact with the Earth's surface, some of it is absorbed by the surface. When this solar radiation is absorbed by the Earth's

	Explanation)	surface, it is heated by conduction. The Earth gives off this heated solar radiation as infrared radiation as heat. (KI Score = 3)
	Critique	The response can be improved by… explaining what kind of energy SR becomes when it is absorbed. (Correct choice for critique)
	Revision	When solar radiation from the sun comes in contact with the Earth's surface, some of it is absorbed by the Earth. The amount of radiation absorbed or reflected depends on the amount of albedo, or ability to reflect solar radiation. This means that an area with high albedo would reflect more solar radiation and an area with low albedo would absorb more solar radiation. When solar radiation is absorbed by the Earth's surface, it is heated by conduction. The Earth gives off this heated solar radiation as infrared radiation or heat. (KI Score = 3)

Missed Opportunities: Turn-Taking and Guessing

Hailey and Tom took turns controlling the computers and answering prompts. Although they engaged with each other and the unit, the nature of their collaboration was primarily strategic in that they alternated responsibility for answering the prompt discussed, but rarely discussed the science content and their understanding of it. They both remained engaged regardless of whose turn it was, but their peer monitoring rarely ventured beyond logistics and accountability with comments such as, "It's your turn," "I'm not going to tell you anything," "You just have to get it better," "You can click there," and so on. Upon encountering an impasse, both partners tended to ask the other to try the step. They neither asked for help nor discussed alternatives. This activity elicited more frequent turn-taking comments such as "Here, you try" that were less commonly observed in other activities. Despite their focus on turn taking, there were instances during the critique and guidance checkpoint steps that provoked moments of content discussion and negotiation. These opportunities were rarely pursued.

The transcript below illustrates one example of a missed opportunity during the guidance checkpoint. Prior to this, they had engaged in frequent turn-taking while attempting revision of the critiqued explanation. However, they did not discuss the critique choices. Tom eventually typed the revision, which consisted of capitalizing one word (Table 6); although Hailey watched attentively, they engaged in an off-topic discussion. When Tom continued to the guidance checkpoint and paused, Hailey asked Tom if he needed help for the first time, but Tom did not take Hailey up on her offer. After multiple instances of turn-taking and failed attempts to pass the checkpoint, they attend to the content of the critique choices for the first time in the activity sequence.

1 Tom: OK, remember "explain that IR comes from the Sun." (T reading previously selected choice; T navigates back to revisit the simulation)
2 Hailey: It doesn't even make sense, though. (T waits for the simulation to load)
3 T: I know. (T begins navigating back to guidance checkpoint without watching simulation)
4 T: This one, right? (T makes a selection)
5 H: I guess? (T scrolls down to hit submit; choice is incorrect)

We see this as an important moment because, in a departure from their usual mode of turn-taking collaboration, Tom asked Hailey to attend to the content of the critique choice (line 1), and Hailey commented on the content to indicate her confusion (2). However, instead of leveraging this opportunity to resolve their dilemma through discussion or by reviewing the simulation, Tom simply agreed with Hailey and navigated back to the checkpoint (3). There, Tom selected another choice without explicating why, but asked Hailey for confirmation (4), who indicated she was not sure (5), which also diverged from their turn-taking mode. However, Tom proceeded to submit his choice without comment. Following this episode, Hailey indicated her frustration and took over the computer. Tom then suggested requesting help for the first time during the activity sequence, saying, "Ask [the researcher], 'cause that is confusing," but neither did so. Eventually, they managed to make a correct guess without discussing the content and proceeded to the next step.

Their written artifacts (Table 6) suggest that the activity design had no impact on Hailey and Tom's progression through the unit. However, by examining their video records, we see moments where the design was successful in provoking opportunities for the dyad to engage in collaborative sensemaking, discussion, negotiation, and reconsideration of ideas, because the guidance disrupted their turn-taking approach to collaboration. Yet, in contrast to Janelle and Ida's case, they rarely capitalized on those moments, proceeding through the activity sequence without discussing or reconsidering their ideas. Further work is necessary to refine the guidance to address these observed limitations so that more students can be supported in making progress.

Table 6. Critique artifacts, critiques, and revisions by Hailey and Tom

Critique Step	Artifact (Assigned Explanation)	Solar radiation was reflected by earth as infrared radiation. (KI Score = 2)

	Critique	The response can be improved by… explaining that IR isn't reflected SR. (Correct choice for critique)
	Revision	Solar Radiation is reflected by Earth as infrared radiation. (KI Score = 2)
Revisio n Step	Artifact (Original Explanation)	It came from the sun and space. (KI Score = 2)
	Critique	The response can be improved by… adding more evidence in general. (Incorrect choice for critique)
	Revision	It came from the sun and space. The heat came from the sun witch is in space. And it will heat the Earths atmosphere and the people. (KI Score = 2)

Summary and Design Implications

Our findings illustrate the value of encouraging students to consider a range of ideas and to capitalize on critique activities. Considering alternatives during critique led to progress in conceptual understanding of scientific phenomena. Although students were assigned explanations of differing complexity in the two conditions, students in both conditions benefited equally from the critique opportunity and were able to make conceptual improvements to their initial explanations during revision, suggesting that both conditions introduced desirable difficulties. The activity sequence designed to provide students with multiple opportunities to consider a range of alternatives was equally successful for critique of explanations that were incomplete and those that included a non-normative idea. The slight trend for critique of the non-normative alternative deserves further study with a larger sample. Since the critique artifacts designed for the study cover a relatively narrow range of complexity, more research is also needed to identify the generalizability of these findings to other critique artifacts, different domains, and to students with different levels of prior knowledge. Future work will also help clarify the specific value of the opportunities to reflect on alternative ideas (such as revising the critique artifact and revisiting content based on guidance) in supporting students to make progress.

By broadening the alternatives for critique and providing multiple opportunities to reconsider ideas, the current investigation showed benefit of critique activities for enhancing students' conceptual understanding. These findings resonate with other investigations of critique activities such as providing critique guidelines (Chang & Linn, 2013). The case studies illustrate the limitations of this approach. While some students were sufficiently prompted by the guidance to reassess their understanding, others needed additional support. A next step is to refine the guidance design to help students such as Tom and Hailey to seriously consider the ideas of their partner.

References

Bjork, R. A., & Linn, M. C. (2006). The science of learning and the learning of science: introducing desirable difficulties. *American Psychological Society Observer*, *19*, 29, 39.

Chang, H.-Y., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching, 50*(7), 858-886.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction*, *24*(4), 467-563.

Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *Journal of the Learning Sciences*, *21*(1), 45–83.

Linn, M. C., Davis, E. A., & Bell, P. (Eds.). (2004). *Internet Environments for Science Education*. Mahwah, NJ: Erlbaum.

Linn, M. C. & Eylon, B.-S. (2006). Science Education: Integrating Views of Learning and Instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology, 2nd edition* (pp.511-544). Mahwah, NJ: Lawrence Erlbaum Associates.

Sato, E., & Linn, M.C. (2011, April). Developing Criteria for Explanations in Science: Scaffolding Peer Critique and Feedback in Technology-Enhanced Instruction. Poster presented at the 2011 American Educational Research Association Annual Meeting, New Orleans, LA.

Shen, J. (2010). Nurturing students' critical knowledge using technology-enhanced scaffolding strategies in science education. *Journal of Science Education and Technology*, *19*(1), 1–12.

Svihla, V., & Linn, M.C. (2012). Distributing Practice: Challenges and Opportunities for Inquiry Learning. In *Proceedings of the 10th International Conference of the Learning Sciences (ICLS2012)* (Vol. 1, pp. 371-378).

Being Mathematical Relations: Dynamic Gestures Support Mathematical Reasoning

Candace Walkington, Southern Methodist University, 3101 University Blvd., Dallas, TX, 75205, cwalkington@smu.edu
Rebecca Boncoddo, Central Connecticut State University, 1615 Stanley Street, New Britain, CT, 06050, boncoddo@ccsu.edu
Caroline Williams, Mitchell J. Nathan, Martha W. Alibali, University of Wisconsin–Madison, 1025 W. Johnson St., Madison, WI, 53706
ccwilliams3@wisc.edu, mnathan@wisc.edu, mwalibali@wisc.edu,
Erica Simon, Southern Methodist University, PO Box 750114, Dallas, TX 75275, ehsimon@smu.edu
Elizabeth Pier, University of Wisconsin-Madison, 1025 W. Johnson St., Madison, WI, 53706, epier@wisc.edu

Abstract: In mathematics classrooms, body-based actions, including gestures, offer an important way for students to *become* mathematical ideas as they engage in mathematical practices. In particular, a type of gesture that we call a *dynamic depictive gesture* allows learners to model and represent fluid transformations of mathematical objects with their bodies. In this paper, we report on two empirical studies – one in which dynamic gestures were observed, and one where these gestures were directed. We conclude that dynamic gestures are a key element in successful justification and proof activities in mathematics.

Introduction

According to the embodied cognition perspective, body-based behaviors associated with intellectual performance are not merely epiphenomenal, but are constitutive of the mental process (Wilson, 2002). One way learning is enhanced is by *grounding* abstract and unfamiliar ideas through situated action and body states (e.g., Barsalou, 2008). Some accounts proffer reciprocity between the action system and cognition such that cognitive states and goals can lead directly (and unconsciously) to actions and actions can induce cognitive states (Nathan et al., under review); in this way, the very boundaries between thinking and acting become blurred. Such embodied perspectives are particularly salient for the domain of mathematics. In effect, one way of *knowing* a mathematical relationship is by *being* the relationship. In particular, learners can enact and therefore *become* mathematical relations is by using *gestures*, an important type of body-based action.

Recent empirical findings lend support to this view. Abrahamson Trninic, and Gutierrez (2012) explored how enactment of the covariation of two constant rates helped foster proportional reasoning. Gerofsky (2011) describes students' accounts of how the ways in which they "become" the Cartesian graphs affects their understanding of the mathematical relations represented. Petrick and Martin (2012) discuss how having high school students physically enact rather than observe geometric relations improved learning gains on conceptual assessment items. As learners engage in the situated practice of mathematical reasoning, body-based actions are an important element of how they become competent members of a community of practice. The ICLS theme, *Learning and Becoming in Practice*, thus connects to our work. Our studies were inspired by observations of teachers and students using gestures to represent dynamic mathematical ideas (Walkington et al., in press), and our research underscores the importance of body-based actions as a way of becoming in mathematical practice.

Here we expand on current research by focusing on a subset of gestures and the utility of these gestures for supporting students' reasoning abilities. Specifically, we explore the distinction between *static* depictive gestures, which display an unmoving, unchanging mathematical object in bodily form, and *dynamic* depictive gestures, which display a mathematical object being transformed using the affordances of the body. We report on two empirical studies designed to explore the nature of mathematical reasoning in the form of proof practices and how these practices are influenced by action. Study 1 uses an observational approach to examine how gestures that occur spontaneously relate to one's justifying and proving. Study 2 uses an experimental paradigm to investigate whether directed gestures can improve proof practices. We explore the implications of these findings for theories of learning and instruction, with a focus on the enactment of mathematical relations.

Theoretical Framework

Embodied Cognition

Theories of embodied cognition posit that cognition is deeply rooted in action and perception (Wilson, 2002). This perspective rejects the view that cognition involves algorithmic processes that use amodal symbol systems, and identifies the body itself as a crucial element in cognition. This implies that mental representations of objects are experiential, perception-based, and multimodal. In mathematics, embodied theories stand in stark contrast to a view of mathematics as an amodal, transcendental, objective feature of the universe. Instead,

embodied theories view mathematics as constructed of body-based experiences of human beings with the world (Lakoff & Nunez, 2000). For example, understanding of number is spatial and tied to bodily orientations (Dehaene, Bossini, & Giraux, 1993), and children approach arithmetic problems using modeling approaches in which they manipulate objects or count with their fingers (Carpenter & Moser, 1984). When learning fractions, actions coupled with interpretations serve as developmental precursors to general mathematical procedures, which can later be enacted mentally (Martin & Schwartz, 2005). Even when working with algebraic equations, students perceive symbols and equations as having concrete, spatial and perceptual qualities (Landy, Brooks, & Smout, 2012). Thus, we posit that all mathematical cognition is embodied. In this work, we study a particular type of body-based action that provides evidence for the embodiment of cognitive processes – gesture.

Gesture

Gestures are an form of action (Goldin-Meadow & Beilock, 2010) that has been theorized to emerge from embodied perceptual and motor simulations that underlie mental imagery and language processing (Hostetter & Alibali, 2008). Gestures can guide attention and communicate spatial, relational, and embodied concepts (Alibali, Nathan, & Fujimori, 2011). Gestures can also serve to link ideas and representations, with gestural catchments (i.e., repeated iconic gestures, see McNeill & Duncan, 2000) creating structural mappings between different entities to show relatedness (Alibali et al., 2011; Nathan, 2008). Recent research has begun to explore how performing gestures can influence the gesturer's thought processes (Goldin-Meadow & Beilock, 2010; Goldin-Meadow, Cook, & Mitchell, 2009). For instance, requiring students to to represent ideas through gesture supports long-term retention of concepts (Cook, Mitchell, & Goldin-Meadow, 2008), and directing students to gesture can instigate the creation of novel ideas (Goldin-Meadow et al., 2009).

One important type of gesture is *depictive* or *iconic* gestures (McNeill, 1992). Here, speakers directly represent objects or ideas with their bodies – e.g., they may form two crossing line segments with their hands, or use their fingers to connect three sides of a triangle. Our research on gesture during mathematical problem solving, as well as research of others (e.g., Goksun et al., 2013), suggests an important distinction between two types of depictive gestures. In *static depictive gestures*, problem-solvers represent an object (like a triangle or line segment), but do not attempt to directly act upon that object. The gesture shows a static representation of a single object that is not interacting with other objects. In *dynamic depictive gestures*, problem-solvers first represent an object, and then engage in fluid transformations of that object using the affordances of their body. For example, a problem-solver might "collapse" a triangle formed with their hands into two line segments on top of each other, or create a rectangle with their hands that "grows" as they move their hands outwards. From an embodied cognition perspective, physical action both results from and initiates cognitive states; thus, performing dynamic gestures with the body might both be a by-product of reasoning processes *and* also give rise to novel ideas. In this paper, we explore the idea of a dynamic gesture, and show how these gestures are important in two studies of students' reasoning when engaging in justification and proof activities in geometry.

Justification and Proof

Justification and proof are challenging practices for students to master as they reach secondary mathematics classes, particularly high school geometry, which more heavily emphasizes this type of complex mathematical thinking (Healy & Hoyles, 2000). Research has shown that students often test examples rather than engaging in general justification (Knuth et al., 2002), and rely on description and perception rather than formal mathematical reasoning (Jones, 2000). Proofs that are mathematically valid have three key characteristics: (1) they are general and show that the argument is true for all possible cases; (2) they involve operational thought with a progression through sub-goals that correctly anticipate the results of mathematical transformations; (3) they involve logical inference in which conclusions are drawn from valid premises (Harel & Sowder, 2005). Harel and Sowder (2005) distinguish such valid proofs from other types of proof by naming them *transformational proofs*. One tool that supports understanding of mathematical proofs with action is Dynamic Geometry Systems. These systems allow students to engage in action-based manipulations of objects on a screen, in order to support students' understanding and exploration of mathematical conjectures (Christou et al., 2004; Marrades & Gutierrez, 2000). Although these technology systems are powerful, we argue that dynamic depictive gestures can provide some of the same affordances, while also remaining highly portable, flexible, and personalized.

Research Purpose

Here we report two studies of learners' engaging in geometric proofs. As we observed learners' proof activities, we discovered that there was a particular class of gestures – dynamic depictive gestures – that seemed important in valid reasoning processes. Thus our overarching research purpose was to explore these dynamic gestures in the context of mathematical conjectures. We examined the dynamic gestures that learners spontaneously produce (Study 1) and we specifically directed learners to use dynamic gestures and tested the effects (Study 2). In Study 1, we identify and describe the characteristics of dynamic gestures and their association with different types of reasoning practices. In Study 2, we seek to tease apart an important distinction regarding whether

dynamic gestures are simply the natural result or by-product of valid mathematical reasoning, or whether they also function to spur novel insights by allowing learners to experience geometric ideas in body-based form. Pen and paper, critical resources typically used by students in mathematics class, were removed in both studies to encourage participants to use their bodies.

Study 1

Our research questions in Study 1 were: (1) What are the characteristics of the gestures that problem-solvers spontaneously produce when justifying a set of geometry conjectures? (2) How are different types of gestures related to production of valid proofs? and (3) How does freedom of gesture production (standing without writing materials vs. seated with writing materials) influence (a) the types of gestures participants produce, and (b) the nature of their mathematical reasoning and proof practices?

Method

Students were solicited to participate in a problem solving experiment on the campus of a selective private university in the South. Fifteen students (9 female, average age = 20.7 years) were asked to provide justifications for 7 geometric conjectures that were mathematically true or false (Table 1). Eleven of the participants had taken Calculus I or higher. Two conditions were alternated among participants. In both conditions, participants were asked to read aloud the conjectures and generate concurrent verbal reports (i.e., think alouds) while being videoed. Participants in the control condition were seated facing the conjectures displayed on screen in front of them and were given a paper and pen. The interviewer sat off to the side but facing the participants and gave only scripted prompts. Participants in the treatment condition were asked to stand within a defined area facing the screen. No paper or pen was provided. The interviewer stood in the same place facing the students and gave the same scripted prompts. Conjectures were presented in random order.

Table 1: Geometric conjectures given to participants in Study 1

<table>
<tr><td>If you double the length and width of a rectangle, then the area is also doubled. (False)</td><td>Given that you know the measure of all three angles of a triangle, there is only one unique triangle that can be formed with these three angle measurements. (False)</td></tr>
<tr><td>The area of a parallelogram is the same as the area of a rectangle with the same length and height. (True)</td><td>The sum of the lengths of two sides of a triangle is always greater than the length of the third side. (True)</td></tr>
<tr><td>The diagonals of a rectangle are always congruent (i.e., they have the same length). (True)</td><td rowspan="2">The segment that joins the midpoints of two sides of any triangle, called the midsegment, is parallel to the third side. (True)</td></tr>
<tr><td>All four-sided figures have angles that add up to 360 degrees. (True)</td></tr>
</table>

Participants' speech and gestures were examined from video. Justifications were analyzed to determine if a participant judged a conjecture to be true or false (*T/F Judgment*). Proofs were analyzed as to whether participants constructed a valid, transformational proof of the conjecture (*Proof*). Gestures were coded into 4 categories: (1) The participant made only static gestures that represented a stationary mathematical object (*Static*), (2) the participant made at least one dynamic gesture that involved a movement-based transformation of a mathematical object (*Dynamic*), (3) the participant drew on their paper and potentially used pointing gestures indicating positions on the paper (*Drawing*), or (4) the participant made no gestures or drawing actions (*None*). The Drawing code could co-occur with Static or Dynamic – for example, the participant may have begun their justification by producing a drawing and gesturing at it, but then abandoned that drawing to engage in standalone Static or Dynamic gestures that were not related to their drawing. Analyses were conducted based on 15 participants generating 7 justifications each ($15 \times 7 = 105$). One justification was missing due to a video malfunction, and another due to a participants' refusal to give a response, for a final count of 103 justifications.

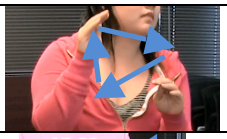
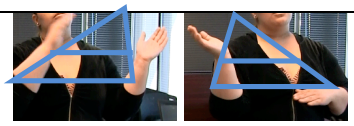
Results

1. Characteristics of Gestures

We noted several different types of depictive gestures that occurred as students provided justifications for the conjectures, which are illustrated in Table 2 below. We coded whether the referent object of the gesture (e.g., the triangle, rectangle, line segment, etc. the gesturer is modeling) was static (i.e., non-moving) or dynamic (i.e., moving). We thus call a gesture that displays a static object to be a static gesture, and a gesture that displays a dynamic object to be a dynamic gesture. This distinction is illustrated with the Static-Trace category in Table 2 – although a trace gesture involves continuous dynamic movement on the part of the gesturer as they outline an object, the *object being depicted itself* is static and non-moving, and thus this was classified as a static gesture. Among the seven conjectures, all had instances of participants using static gestures and drawing, while six of the

seven had instances of participants using dynamic gestures. Of the 15 participants, 11 used dynamic gestures at some point while justifying a conjecture, and 14 used static gestures.

Table 2: Types of depictive gestures observed

Referent Object is…	Gesture Type	Description	Image
Static	*Trace*	Participant traces over the outline of a stationary line or a shape in the air, similar to drawing on a page. *(On right, participant traces triangle in the air)*	
Static	*Represent*	Participant uses hands or fingers to physically formulate a complete or semi-complete object. *(On right, participant forms a right triangle)*	
Dynamic	*Rotate/ Reflect/ Translate*	Participant "picks up" an object represented with their hands/fingers, and then slides, rotates or reflects to change its orientation or position. *(On right, participant makes a parallelogram with hands, and then slides the two consecutive angles together to show they equal 180°)*	
Dynamic	*Dilate*	Participant begins by representing a static object with hands, and then moves hands outwards and inwards to show the object growing or shrinking. *(On right, participant makes the triangle he forms larger then smaller)*	
Dynamic	*Test Interactivity*	Participant modifies one element of an object to predict impact on the rest of the object. *(On right, participant shifts apex of triangle to see what happens to midsegment)*	

2. Relationship between Gestures and Valid Proofs

We hypothesized that the dynamic gestures in particular were an important way to engage with mathematical relationships during justification practices and to formulate transformational proofs. Indeed, this is the premise behind dynamic geometry systems like Geometer's Sketchpad and Cabri Geometry. Although this was a small data set, we looked at trends in the relationship between gesture types and valid proofs to see if this idea was worthy of further investigation. Results are shown in Table 3. Dynamic gestures were associated with the highest accuracy on both true/false judgments and production of transformational proof. Making no gestures or drawings was associated with the lowest performance. Static gestures and drawing actions fell in between.

Table 3: Associations between gesture codes and average accuracy, for each geometric conjecture ($N = 103$)

Gesture Performed	% Correct on T/F Judgment	% Formulating Valid Proof
Dynamic	90.9%	63.6%
Static	74.3%	34.3%
Drawing	84.4%	27.3%
None	57.7%	11.5%

Table 4: Associations between condition, gestures, and average accuracy, for each conjecture ($N = 103$)

Condition	Accuracy on T/F	Accuracy on Proof	% of justifications involving dynamic gesture(s)	% of justifications involving only static gestures
Pen	76.4%	20.0%	12.7%	30.9%
No Pen	75.5%	52.1%	30.6%	36.7%

3. Relationship between Condition (Pen/No Pen) and Valid Proofs

Given that drawing gestures were associated with lower accuracy on formulating valid proofs than depictive gestures (static and dynamic), we also investigated how students' accuracy and tendency to gesture varied by

whether or not they were given a pen and paper. Results are shown in Table 4. Participants in the "No Pen" condition had higher accuracy when formulating proofs, and were more likely to make dynamic gestures.

Discussion

This study suggests that dynamic gestures are an important component of formulating and communicating valid proofs in geometry. Dynamic gestures may be promoted when learners are denied tools of pen and paper. These traditional tools may in some cases be less productive for mathematical reasoning than simply encouraging students to use their bodies. This study also suggests some characteristics of dynamic gestures used in geometry, and illustrates different types of dynamic gestures used across a variety of conjectures. Dynamic gestures appear to be an important part of justification and proof activities, so their use should be encouraged by providing students with greater freedom to gesture. In Study 2, we investigate the potential of dynamic gestures further in an experimental paradigm, by explicitly directing participants to perform dynamic gestures prior to a proof task.

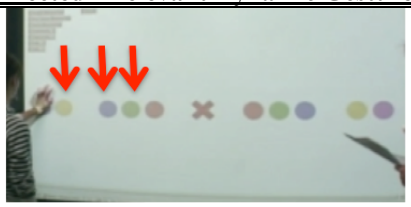
Study 2

Our research questions in Study 2 were: (1) Does explicitly directing participants to perform dynamic depictive gestures influence their accuracy when justifying a geometry conjecture? (2) Is there a significant association between learners generating their *own* dynamic gestures during the process of proof and justification, and their accuracy when justifying a geometry conjecture?

Method

Participants were 80 undergraduates (44 female, average age = 19.5 years) enrolled in a Psychology course at a large Midwestern university. Sixty of the participants had taken Calculus I or higher, and their average ACT/SAT math percentile was 87. Participants provided a justification for the triangle conjecture (*"For any triangle, the sum of the lengths of any two sides must be greater than the length of the remaining side."*) while being video-recorded. They key idea was that if the two sides were shorter they would not be able to reach the endpoints of the remaining side to close the triangle. Of the 80 participants, 40 were first explicitly directed to perform relevant dynamic gestures related to the triangle conjecture (Table 5), and were directed to use their bodies to form changing versions of the referent object (a triangle). In effect, they formed possible and impossible triangles with their arms or hands, with one side of the triangle dynamically "growing" until a triangle could no longer be formed. They were not told these gestures were related to the conjecture. The other 40 were directed to enact irrelevant gestures (Table 5). Of the 40 participants in the relevant condition, only 4 reported that they saw a connection between the directed gestures and the conjecture, and the results were similar with or without those participants. As Study 1 suggested dynamic gestures were facilitated in absence of pen and paper, no participants were allowed to use these tools. Participants were asked to think aloud.

Table 5: Dynamic gestures that participants were directed to perform

Directed Relevant Dynamic Gestures	**Directed Irrelevant Dynamic Gestures**
	
Participants touched concentric pairs of circles with outstretched arms, with the last pair being too far apart to touch. (A second version involved hands instead of arms, not shown here.) In pilot work, we found participants spontaneously using gestures similar to these.	Participants were asked to walk back and forth in front of concentric pairs of circles, touching one circle at a time. (A second version involved smaller circles that were closer together, not shown here.)

Analyzing video records, we coded each participant's judgment of the conjecture as true or false (*T/F Judgment*), and coded whether they generated a valid transformational proof (*Proof*). A kappa reliability of 0.82 was obtained by 3 coders on a list of proof categories adapted from Harel and Sowder (2005) that also included the T/F judgment. We then coded the gestures participants spontaneously made while justifying the conjecture. Gestures were coded into 3 categories: (1) The participant made no gestures (*None*), (2) The participant made only static gestures that represented a stationary triangle or triangle part (*Static*), and (3) The participant made at least one dynamic gesture that involved a movement-based transformation of a triangle (*Dynamic*). This coding represented only self-initiated gestures that participants produced while justifying the conjecture, not gestures that participants were directed to perform; inter-coder reliability was kappa = 0.78. We fit logistic regression

models predicting a correct T/F Judgment or Proof (coded as 0/1). Predictor 1 was whether the participant had been directed to perform relevant dynamic gestures prior to being given the conjecture (Condition-Experimental) or whether they had been directed to perform irrelevant gestures (Condition-Control). Predictor 2 was whether the gestures participants spontaneously produced during justification were Not Dynamic or Dynamic. We controlled for math achievement using self-report ACT/SAT math percentile as a covariate.

Results

1. Effects of Dynamic Gestures that are Explicitly Directed

We first examined whether the relevant dynamic gestures we directed participants to perform prior to being given the triangle conjecture affected their subsequent accuracy when justifying the conjecture. Although participants who were directed to perform relevant gestures constructed a valid proof of the conjecture more often than participants who were directed to perform irrelevant gestures (50% vs. 40%), this difference did not reach significance ($z = 0.095$, $p = 0.924$). Likewise, participants who were directed to perform relevant gestures correctly judged the conjecture was true more often than participants who were directed to perform irrelevant gestures (92.5 vs. 82.5% of cases), but this difference also did not reach significance ($z = 1.317$, $p = 0.188$).

At the end of the session, all participants who had performed directed relevant gestures ($n = 40$, first column of Table 5) were informed of the relevance of these gestures to the triangle conjecture. During this debriefing these participants were then given an opportunity to provide a second justification for the triangle conjecture based on this information. The 40 participants were more likely to give a correct proof on this second attempt ($z = 2.190$, $p = 0.0285$), with their chances of obtaining a correct proof increasing from 50% to 70%. Directed relevant gestures may only be effective when the learner is explicitly made aware of their relevance.

2. The Impact of Dynamic Gestures produced during Justification and Proof Activities

Although all participants were explicitly directed to perform some gestures prior to being shown the triangle conjecture (Table 5), we were also interested in the gestures that participants spontaneously produced as they engaged in the proof and justification activities. Figure 1 gives an example of a dynamic gesture sequence that was spontaneously used to prove the triangle conjecture. The participant gives a specific example of an equilateral triangle and then generalizes to all triangles. She moves her thumbs apart, representing the bottom of the triangle as she says that "it wouldn't connect," and then moves her thumbs together as she says "make 'em a triangle." Finally, she collapses the triangle into a line as she says, "they would be like flattened out." This participant produces gestures where she is dynamically modifying one aspect of the triangle (the side lengths) using her body, and then using gestures to see what would happen to the rest of the triangle as a result.

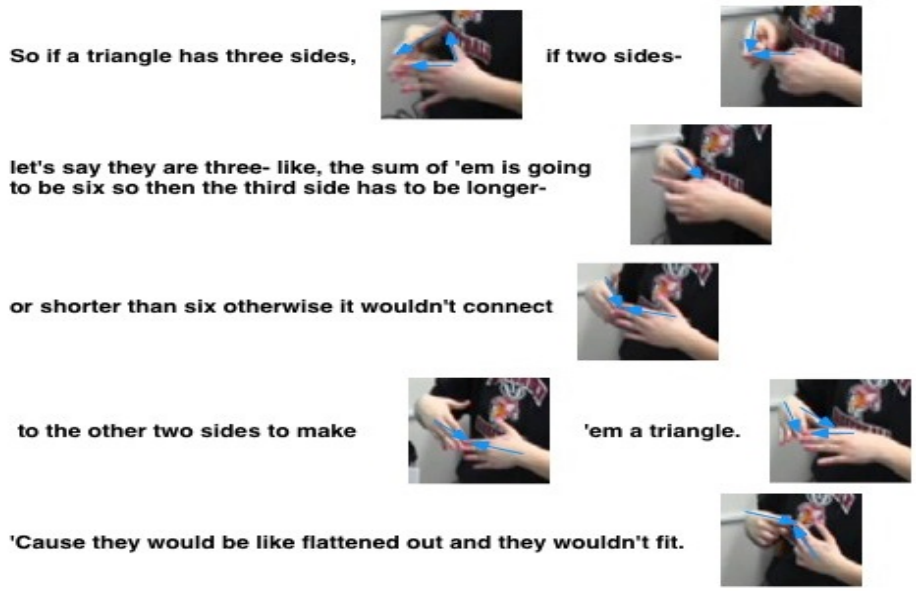


Figure 1. Participant uses dynamic representation gestures to justify triangle conjecture

Many participants in both conditions made their own spontaneous dynamic gestures, and we were interested in whether these gestures were associated with correct T/F judgments and proofs. Participants who produced dynamic gestures were more likely to correctly judge the conjecture to be true than participants who did not produce dynamic gestures (100% vs. 75%). This trend was similar across conditions (experimental/control). However, this difference did not reach significance in the models due to the error term

associated with the ceiling effect. Participants who spontaneously produced dynamic gestures were also more likley to provide a correct justification than participants who did not spontaneously produce dynamic gestures (57.5% vs. 32.5%). This difference was significant in the regression model ($z = 2.90$, $p = 0.00375$, $d = 1.0$). However, being directed to perform relevant dynamic gestures prior to being given the conjecture did not influence whether participants actually used dynamic gestures to justify the conjecture. In both condictions, exactly 12 of the 40 participants (30% of participants) made spontaneous dynamic gestures.

Discussion

Study 2 highlights the importance of learners' generating their own dynamic gestures as they spontaneously engage in justification and proof activities. These spontaneous gestures may simply be outward evidence of a stronger understanding of the geometric relations, or they may also indicate that giving dynamic imagery body-based form can support and enhance mathematical reasoning. In other words, spontaneous dynamic gestures may simply be a by-product that is often seen with advanced reasoning, or alternatively students may actually learn from "being" the geometric shape as they generate spontaneous gestures. Given episodes like Figure 1, in which a participant seems to be actively attending to and experimenting with geometric ideas using her body, we believe the latter is a strong possibility. However, Study 2 also cautions that directing participants to perform the physical action of dynamic gestures may not, in isolation, be useful to formulating valid proofs. If dynamic gestures are directed, the learner may need to pay explicit attention to and reflect on the gestures for them to be useful. Asking the learning to "become" a triangle may be of limited usefulness to proof generation if the learner believes they are simply making meaningless, non-mathematical motions. Thus, giving students prompts that relate directed dynamic gestures to the task at hand is important to facilitating valid proof generation.

General Discussion and Conclusions

The current studies support the idea that the dynamic gestures that learners spontaneously produce allow learners to utilize the affordances of their body to ground their understanding of mathematical relationships. Thus, the findings suggest that complex reasoning in a challenging area of mathematics can be fostered by recruiting body-based resources. Study 1 showed that participants who spontaneously used dynamic gestures demonstrated superior mathematical reasoning. The findings suggest that when people enact the key mathematical relations of a task in dynamic, body-based form, they are better able to accurately assess the validity of mathematical conjectures and are more likely to generate valid mathematical proofs to warrant their judgments. Study 2 reinforced the important relation between dynamic gestures and valid proofs, but also suggested that those who were directed to enact relevant relationships through dynamic gestures were more likely to construct valid proofs, provided the purpose of the directed gestures was made explicit.

Although further study is needed to clarify these findings, the studies corroborate the view that reasoning through enactment is associated with conceptual development in an area of study that students find quite challenging. In a recent study, Goksun et al. (2013) found that the gestures of adults with high spatial reasoning abilities were more likely to contain dynamic information about mental rotation. Similarly, Ehrlich et al. (2006) found that children who produce dynamic gestures performed better on a mental rotation task. Our findings further suggest that behaviors such as writing may inhibit dynamic gesture production and impair mathematical reasoning. In similar fashion, Martinez (2012) found that participants who had to type their responses to a science test made fewer relevant inferences than participants who instead spoke their responses and were consequently free to gesture. Both that study and the current one raise questions about whether assessment practices that impinge on people's ability to freely produce gestures may have the unintended consequence of impairing their abilities to generate inferences. Dynamic gestures offer a unique way for learners to "become" a mathematical idea as they engage in learning mathematical practices. Our work seeks to make explicit the importance of identifying, analyzing, and facilitating the dynamic gestures in the classroom.

Acknowledgements

Thank you to Alyssa Holland, Dan Klopp, and Jessica Waala for their contributions. This work was made possible through grants from the National Science Foundation (DRL-0816406), the Institute of Education Sciences (R305B100007), and the Department of Teaching and Learning at Southern Methodist University.

References

Abrahamson, D., Trninic, D., & Gutiérrez, J. F. (2012). You made it! From action to object in guided embodied interaction design. In P. Freebody, T. de Jong, E. Kyza, & P. Reimann (Eds.), *Proceedings of the International Conference of the Learning Sciences: Future of Learning*: University of Sydney.

Alibali, M. W., Nathan, M. J. & Fujimori, Y. (2011). Gestures in the classroom: What's the point? In N. L. Stein & S. W. Raudenbush (Eds.), *Developmental Cognitive Science Goes to School* (pp. 219-234). New York: Routeledge.

Carpenter, T., & Moser, J. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, *15*(3), 179-202.

Christou, C., Mousoulides, N., Pittalis, M., & Pitta-Pantazi, D. (2004). Proofs through exploration in dynamic geometry environments. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 215-222). Bergen, Norway.

Cook, S., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047-1058.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371-396.

Ehrlich SB, Levine SC, Goldin-Meadow S (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology*, *42*, 1259–1268

Gerofsky, S. (2011). Seeing the graph vs. being the graph: Gesture, engagement and awareness in school mathematics. In G. Stam & M. Ishino (Eds.), *Ingrating gestures* (pp. 245-256). Amsterdam: John Benjamins.

Göksun, T., Goldin-Meadow, S., Newcombe, N., & Shipley, T. (2013). Individual differences in mental rotation: what does gesture tell us?. *Cognitive processing*, *14*(2), 153-162.

Goldin-Meadow, S., & Beilock, S. (2010). Action's influence on thought: The case of gesture. *Perspectives in Psychological Science*, *5*(6), 664-674.

Goldin-Meadow, S., Cook, S., & Mitchell, Z. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*(3), 267-272.

Harel, G., & Sowder, L. (2005). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, NCTM.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396-428.

Hotstetter, A., & Alibali, M. (2008). Visible embodiment: Gestures as simulated action. *Psychometric Bulletin & Review*, *15*(3), 495-514.

Knuth, E., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In S. Mewborn, P. Sztajn, D. Y. White, H.G. Wiegel, R.L. Bryant, K. Nooney (Eds.), *Proceedings of the 24th Meeting of PME-NA, Vol. 4*, (pp. 1693-1700). Athens, GA.

Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics*, *44*(1/2), 55-85.

Lakoff, G., & Nunez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Landy, D., Brookes, D., & Smout, R. (2012). Modeling abstract numeric relations using concrete notations. *The 33rd Annual Conference of the Cognitive Science Society*, Boston, MA.

Marrades, R., & Gutierrez, A. (2000). Proofs produced by secondary schools students learning geometry in a dynamic computer environment. *Educational Studies in Mathematics*, *44*, 88-125.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*(4), 587-625.

Petrick, C., & Martin, T. (2012). Learning Mathematics: You're It vs. It's It. In P. Freebody, T. de Jong, E. Kyza, & P. Reimann (Eds.), *Proceedings of the International Conference of the Learning Sciences: Future of Learning* (ICLS 2012). Sydney: University of Sydney / ISLS.

Martinez, C. V. (2012, August). Manipulations of Gesture Production Indicate the Role of the Body in Situation Model Development: Implications for Computer-based Assessment. Dissertation to the University of Wisconsin-Madison. (ProQuest Publication No. 3522569).

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: Chicago University Press.

McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.). *Language and gesture: Window into thought and action* (pp. 141-161). Cambridge: Cambridge University Press.

Nathan, M. J. (2008). An embodied cognition perspective on symbols, grounding, and instructional gesture. In DeVega, M., Glenberg, A, M. & Graesser, A. C. (Eds.) *Symbols, Embodiment and Meaning: A Debate* (pp. 375-396). Oxford, England: Oxford University Press.

Nathan, M., Walkington, C., Boncoddo, R., Pier, E., Williams, C., & Alibali, M. (under review). Actions Speak Louder with Words: The Roles of Action and Pedagogical Language for Grounding Mathematical Reasoning.

Walkington, C., Nathan, M., Wolfgram, M., Alibali, M., & Srisurichan, R. (in press). Bridges and barriers to constructing conceptual cohesion across modalities and temporalities: Challenges of STEM integration in the precollege engineering classroom. In J. Strobel, S. Purzer, & M. Cardella, (Eds.), *Engineering in Pre-College Settings: Synthesizing Research, Policy, and Practices*. Purdue University Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9,* 625-636.

"Learning to Live": Expansive Learning and Mo(ve)ments Beyond 'Gang Exit'

Line Lerche Mørck, Aarhus University, Tuborgvej 164, DK-2400 Copenhagen, llm@dpu.dk

Abstract: The paper develops a social practice theoretical framework to analyze expansive learning in relation to Danish gang exit intervention. We follow 22-year-old Bilal's movements and moments in and out of gang communities, and how Bilal, collectively with his mentor and teacher Jesper and other role models is 'walking and expanding the margins' of Danish exit intervention. At age 12, Bilal started a trajectory becoming 'more of' and later 'less of' a criminal gang member. The paper explores new belongings and (lack of) meanings in and across educations, interventions and gang communities, and how Bilal through recognition in his new communities reinvents himself, us and others, moving beyond marginal positions in and across contexts such as Danish people college, the prison, as well as other intervention contexts organized in a collaboration between his local municipality, Jesper and communities with students, professionals and former gang members.

Introduction

What kind of analytical framework and methodology is relevant to explore and understand learning and life processes when young men change their lives and move beyond gang life? This paper proposes a social practice theoretical framework to explore the young men's movements along very complex learning and life trajectories, as part of social practice research, which is not bound to certain (single) intervention contexts. This framework and methodology explores moments and movements beyond gang life as expansive learning in and across many different communities and action contexts (including various intervention contexts).

The Danish 'Gang Exit' Problematic: Enduring Struggles and Contentious Practice

In Denmark, major changes in gang cultures and intervention practices have been observed recently. After a 'drive-by' killing in 2008 in a socially deprived area of Copenhagen, we observed a sharp rise in shootings and gang conflicts involving new subgroups to the biker gangs and new geographically street socialized groups beginning to adopt more gang identities, name themselves, and use guns to protect their territories in conflicts with each other. The conflicts have escalated into what some of the involved groups call "gang war," involving biker gangs, their subgroups, and different street gangs as well as geographically based street communities. Over the last couple of years, the conflict has become more complex, involving more and more parties, each with several fronts to fight and protect (Mørck et al., 2013). The rise in cultural activities labeled "gang activities" is paralleled by a rise in intervention initiatives. In April 2011 the Danish Ministry of Justice launched a "framework for gang exit programs" (Justitsministeriet, 2011, own translation). Since then, we have seen an increasing awareness in Denmark of the importance of targeted interventions aimed at those looking for assistance in leaving a criminal group, a street gang or a biker gang. Several so-called "gang exit initiatives" have emerged, organized by the state in prison and probation services, by the municipalities, new NGOs and private organizations. The societal goal of these exit units, programs and/or interventions is to support young men like Bilal in moving beyond gang life, leaving criminality and violence behind. There is still no published research in these new exit programs in Denmark, but many of the initiatives, especially the state initiatives by police, prison and probation services, have been criticized by (former) gang members and professionals, at 'gang seminars' (3), in the media and in ongoing (not yet published) research. One critique is the dilemma of police engagement in the exit initiatives, which tends to reproduce distrust. But there are also alternatives to these state established exit programs: Some mentors and project leaders, like Jesper and Ali, are working as part of various contexts, including people's colleges, prisons, NGO's, and in collaboration with municipalities, trying to expand possibilities and transcend learning barriers for young men who are trying to move beyond gang life. The NGO alternatives to exit intervention are also disputed in the Danish media, especially if they involve mentors, cultural activities and communities which include members with a personal criminal history.

Social Practice Theory of Expansive Learning as Movements Beyond Marginalization

Figure 1 illustrates the cornerstones of this analytical framework of expansive learning as the following: Practice, Participation, Meaning and Social Self-understanding (Mørck 2007; 2010; 2011). Individual and collective subjects (Nissen, 2012) learn through participation in practice, in and across action contexts (Nissen, 2012; Dreier, 2008). Critical psychology conceptualizes subjective meaning; emotions should be understood as

closely connected to cognition and agency, and emotions help us to sense our subjective being and what is meaningful:

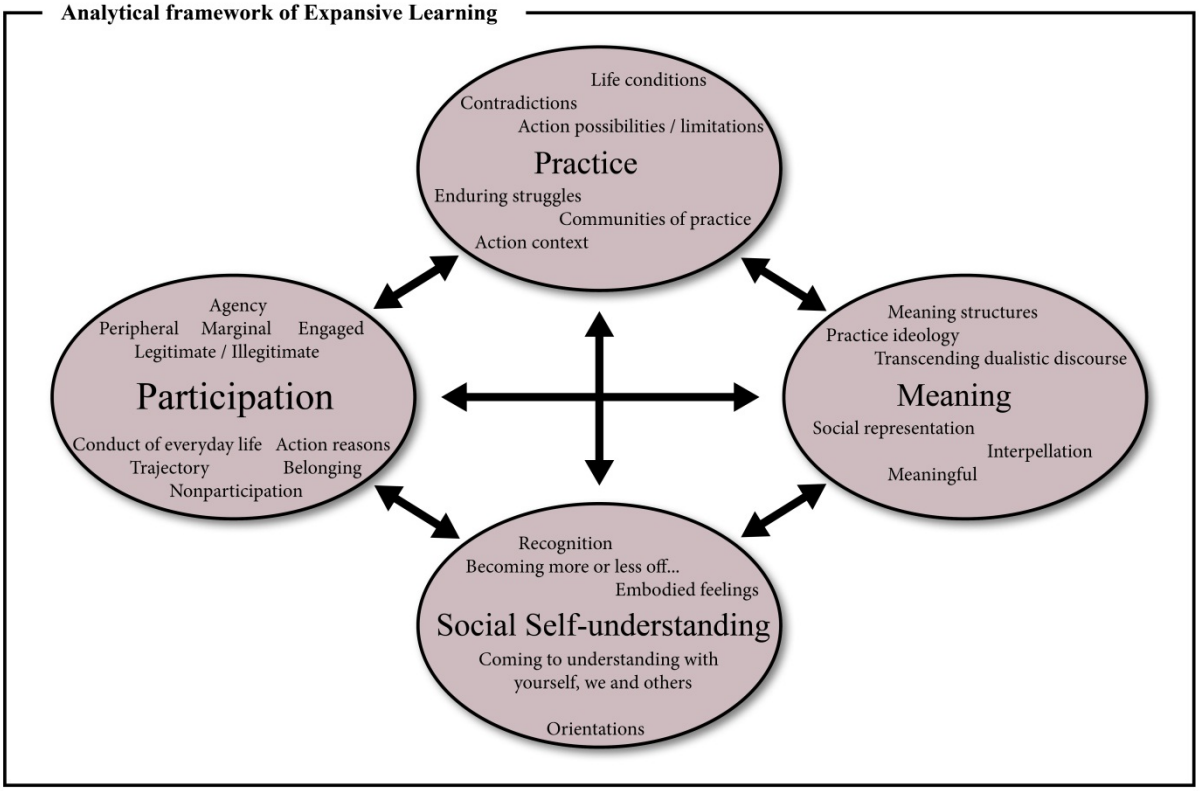


Figure 1. Analytical framework of Expansive learning

"Meaning structures," or culture, (Holzkamp, 2013b, p. 278) are understood as possibilities to act, dialectically related to both objective societal conditions and the subject's action reasons (Holzkamp, 2013b, p. 285). Culture or meaning structures are mediated by historical, economic, political, institutionalized arrangements as part of enduring struggles (Holland & Lave, 2009), and as discourses, they are continually (re)produced in local practice ideologies as part of communities of practice (Mørck, 2010) affecting the young men's social self-understanding (Holzkamp, 2013b; Kristensen & Mørck, 2014), their movements and change in orientations, and belongings in and across communities of practice (Wenger, 1998).

The concept of expansive learning (Kristensen & Mørck, 2014) is inspired by Holzkamp (2013a), who differentiates between "expansive learning," which is meaningful and in line with one's life interests, and "defensive learning" where the learner experiences being "cut off from the joint control over the living conditions, thrown back on [one]self, controlled by immediate threats and needs" (ibid, p. 124). Mørck (2010) further develops Holzkamp's notion of defensive and expansive learning, analyzing additionally how contradictions may be transcended as part of collective agency, with a special focus on mechanisms involved in reproduction of marginalization and/or processes of partly transcending and moving beyond marginal positions.

Being a (former) gang member is a marginal position, which involves enduring struggles, conflicts, tensions and contradictions (Holland & Lave, 2009). Movements beyond demonization and gang labeling involve enduring struggles with dualistic discourses of good and evil. Movements beyond also include a notion of telos of participants' learning and how the telos, directionality, or orientations change over time (Lave, 1997). Expansive learning includes complex processes of contradictory meaning-making and interpellation: "Interpellation is when the subject recognizes herself as recognized in this unique but universal identity, and with this responsibility given as meaningfulness defined in the ideology" (Nissen, 2012, p. 193).

Participatory Social Praxis Research and Mo(ve)ment Methodology

This paper draws on history-in-person ethnography (Lave, 2011; Holland & Lave, 2009) and social practice research (Mørck & Nissen, 2005), developing a combination of intervention practice analyses combined with an embodied mo(ve)ment methodology inspired partially by collective biography. Like Davies & Gannon (2006), this mo(ve)ment field work methodology explores embodied moments and movements of importance, but the methodology differs from collective biography work in that it is ethnographic, combining multiple empirical sources (Kristensen & Mørck, 2014), which are reflected in a collective mo(ve)ment interview and later in co-

researcher feedback on the analysis. The mo(ve)ment interview has several parallels to Holland's in-depth "identity trajectory interviews" (Holland & Lave, 2009, p. 7); it is ethnographic, theoretically it explores senses of self and others as meaning-making across a life span, it explores movements, moments, struggles and tensions as dialectics between collective and individual movements, producing continuation and change in the conduct of everyday life (Dreier, 2008, Osterkamp, 2001).

I have followed the gang seminars organized by Jesper, Ali, myself and others since 2009, researching it as a boundary community (Mørck et al, 2013). With this paper I move more in depth, trying to understand the embodied feelings, motivation, agency and complex struggles, problems and expansive possibilities of these moments and movements. I interviewed Bilal and Jesper together in relation to their perspectives on mo(ve)ments of importance in each of their lives. In preparation for the interviews, I watched a documentary that some of the co-students at the People's College had produced about Bilal and Jesper with the theme, "How People's Colleges make a difference." In the documentary, co-students had interviewed Bilal, Jesper, other students and teachers, and the school principal. I also observed two sessions in the course about 'gang criminality' that Bilal and Jesper were teaching together, and I read a book manuscript that Jesper had written about his criminal past and present life. From all these empirical sources, I produced an interview guide, where I shared my preliminary analysis of what I thought might be moments and movements of importance in Bilal's life conduct. After in-depth discussion of these, there was an open time, where they could introduce and reflect on other moments. Later, field notes were taken when my colleagues and I attended more presentations about changes in Bilal's life and dialogues of interventions at a conference and a seminar. Finally, I watched a second video about new changes in Bilal's life and participated in other local cultural activities organized by Bilal, Jesper and other People's College students, where I met and talked to family members. In and across all these contexts, I explored their participation and change in positions, their expanded agency, meaning-making processes, and their representation of themselves, their communities and other parties, to a variety of people and audiences.

Ethics, Dilemmas and Contradictions

This social practice research approach shares ethics and participatory goals with social justice research projects: 1) humanizing marginalized subjects, 2) contributing to social reform and social justice, and 3) creating conditions for a dialogical relationship between researchers and research participants (Brotherton & Barrios, 2004). As Michelle Fine describes, the ideal of social justice projects "marks a space of analysis in which the motives, consciousness, politics, and stances of informants and researchers/writers are rendered contradictory, problematic, and filled with transgressive possibilities" (Fine, 1998, p. 141).

We seek to build the practice research on common interests (Mørck & Huniche, 2006, p. 7) as part of joint ventures; we exchange and participate from very different positions in and across different communities. I participate from multiple positions, as a person, activist and researcher (Khawaja & Mørck, 2009), as researcher and presenter, we also co-organize seminars together, as well as contribute to community building and knowledge exchange in our networks. Jesper and Bilal are invited into the work as co-researchers (Nissen, 2000; Mørck, 2000), reading and commenting on papers like this before publication.

Due to this very sensitive field (Jacobsen & Kristiansen, 2001) and the complexity in participating from multiple positions, this research also presents difficult ethical struggles and contradictions (Fine, 1998). Some of the persons I have interviewed and observed may be under surveillance by police, rival gangs and/or gangs with whom they were previously affiliated. Additionally, as researcher, I will also from time to time find myself in conflictual contradictions and dilemmas, between institutionalized ethics (such as anonymity) and social justice ethics (Fine, 1998). It is a dilemma that Jesper is a public figure in the media debate, because it challenges the ethical standard of anonymity bringing about contradictions in our joint venture when I still attempt to conceal Bilal's identity.

Because of this dilemma, my researcher team and I continually discuss and struggle to produce new standards of ethical social practice, which both expand institutionalized ethics, e.g. consent practices: here the informants give ongoing consent and permission regarding the use of data in articles and papers. I work closely with gatekeepers such as Jesper to be able to practice social justice ethics in a way that partly transcends the ethical dilemmas. Jesper continually helps consider dangers and analyze empirical data for sensitivities in specific times and contexts of presentation. By doing very important social justice work in this field, he also reproduces trust (Jefferson 2004, p. 40) in relation to the other co-researchers, such as Bilal and Ali.

Findings and Empirical Analysis

To reach an in-depth understanding of Bilal and his collective remembered embodied moments and movements, I attempt simultaneously to be informed by the described social practice analytical framework of expansive learning and to analyze with open-ended curiosity. Here, I present three main findings as forces of special importance for Bilal's movement beyond gang life: 1) Motivational aspects very slowly take form over time as embodied feelings of lack of meanings and new experiences of what is meaningful, 2) Longing for belonging

reflects a process of opening up, building trust and new friendships, 3) Recognition is produced through expansive interpellations that slowly and over time give substance to Bilal as a unique, legitimate, and productive member of new communities.

Embodied Feelings of (Lack of) Meaning

> This is one of the lessons that meant the most to me here at the people's college. Ali mentions that about coming to People's College and learning to live. […] For some reason, I also wanted to get up and leave the room, because I felt that even though Ali sat and talked about himself, it went in and hit me so hard personally. Ali puts feelings into words that I myself have not been able to express.

Bilal describes a strong embodied feeling from Ali's presentation that "went in and hit him so hard" that he felt like getting up and leaving the room. But he also describes how he, at that moment, became much wiser about himself by listening to Ali's reflections of how he changed his life through becoming a People's College student and an active part of the community and life there. For a very long time, Bilal had struggled to find words for what he was going through; he "had not been able to see it." But listening to Ali, he realized that he had "started to live life as normal people do," and knew this was a tremendous change in his life. He sat back and thought, "When did I ever live like this before?" and realized that it was when he was a little boy, before he was 12 years old.

Bilal has participated from marginalized positions most of his life, in and across different contexts. Back in primary school, he spent most of his time "being punished, on the black bench." He recalls being placed outside the classroom four out of six hours in school. At 12, his life was already a daily struggle; he was involved in many fights with other children in the school yard, he often had problems doing his homework, and the teachers responded by writing complains to his parents, who's response was to discipline Bilal physically at home.

One day, at the age of 14, he stabbed another kid in the school yard with a knife. He describes how he suddenly sensed a new kind of power: others being afraid of him. He was then forcibly removed from home by the authorities and placed at a special boarding school, where he became a member of a new community of older boys who were much more criminally experienced than Bilal. Two years later, when he got back to his home town at the age of 16, he felt "well educated in criminality." One day, several years later, he was arrested while carrying a weapon and a changed life in prison began. He describes the time in prison as feeling "very lonely," and he felt "let down by the system" and "let down by his brothers." He spent about 8 month in isolation; he tried to get into the prison and probation exit program, but got rejected and didn't know why. He felt very disillusioned over his time in prison:

> I think I was the kind of person, who did not care much. Once you have done time in solitary confinement for so long, you have nothing left, neither for the system nor anything. Of course, this one social worker was really kind, but I felt let down by the system, so I had no great enthusiasm at that time.

He did, though, express a little surprise when his social worker from the municipality visited him in prison, because she bothered to travel all the way across the country just to visit him.

Shortly after the visit, he was released and found himself at a release party with his "own brothers." It was a big party; some of the other brothers were also released, and there were many tables with beers, drinks and ladies, with all in all around 250 people gathered. Bilal describes the 12 minutes he lasted at the party:

> It was a really, really unpleasant experience, and I had to get in the car, and then I got one of my friends to drive me home. [...] I have never felt so surrounded by people, so pressed, claustrophobic, you are about to throw up and break a sweat and can't help it, you look around and it all stutters in slow motion, as if you were on drugs.[...] I could not acclimate to the many people at the club. I had ended up with a lot of hatred for many of my own brothers. [...] Then it did not fucking matter that they adored you, or that I could get a new status. I went in for something that I believed in - and I later found out, that it was not like that. This I cannot fight for. That was the position I had in the beginning [at the release party].

He also experienced nausea, the few other times he visited the gang environment. The gang life is slowly losing its meaning; he can sense it as embodied affects, but it was only after he heard Ali's presentation that he started to be able to put the meaning-making processes into words. Bilal's meaning-making, including sensations of lack of meanings is described as bodily sensations "hitting [him] hard, personally." It involves

struggles and contradictions, for instance wanting to leave the room even though Ali's presentation is very meaningful to him and his social self-understanding, a feeling nausea even in the presence of drinks, ladies and people that adore him. Concrete action reasons are in some way diffuse, like action reasons for the sensations of nausea or hatred, but we learn that these sensations were important and constellated as either very meaningful or signifying a lack of meaning that pulled him in different directions in life. The action reasons are mainly described in a language of embodied affects: Feelings of being let down, nausea, hatred, or being pressed and claustrophobic, and the sense of lack of meaning, something he no longer can fight for.

Longing for Belonging: Processes of Opening Up, Building Up Trust and Friendships

Starting at the People's College, Bilal hoped for a "fresh start" and he had no plans to tell his own story to the other students. He was aware that most people would find it very difficult to understand his situation. Personal story telling as a (former) gang member entailed a risk that other students would label him as a bad person and building up walls between them. Bilal had a deep fear of being lonely or isolated with no network or friends, which also made him hesitate in the process of opening up and moving beyond a mute positioning. As analyzed below, many moments were important steps for the process of opening up, building trust and friendships, and, in the beginning at the People's College, Bilal was still affiliated with the gang:

> I still have one leg in the [gang community]. The thing I fear the most, is to lose my social network. I have one foot in each camp. I am uncertain that this is the way I want to go. I want the people at school to get to know me as a person, and I don't want them to judge me from my history.

The first couple of weeks, Bilal also had difficulties trusting his mentor and teacher Jesper, but this changed:

> I just had a feeling [...] and I asked him, and he answered me honestly and it meant that I could let my guard down and think; 'okay, I am not the only one here, there is also another person who understands my situation better than anyone else'

At this time, Jesper had begun planning to tell his story to his colleagues. He had only recently told the principal openly about his criminal and radical left wing past. Jesper was writing a book and showed Bilal the foreword. To tell about his past involved worries; his parents had been worried that he could risk losing his job. Bilal followed Jesper closely in this process of opening up, showing solidarity by asking about the responses of the colleagues and the principal.

The change from "mentor-mentee-ping pong" to more mutuality and more trust accelerated when Bilal and Jesper started kickboxing one hour every morning. Bilal had trouble sleeping at night. And the first months at the People's College, he had a very hard time getting up in the mornings. He was late for 'the daily morning song' and often also late for the first lesson of the day. But after Bilal started to coach Jesper in kick boxing, Bilal and Jesper started a friendship, with more mutual talks during breakfast. They both opened up to more sensitive issues, such as crises in love relations. Bilal expressed that this was a big change; he was not used to talking about family relations and love relations in his other communities:

> It was a great feeling, a liberating feeling to let my feelings out, but I know - at that time if there is something that can hit you again, it's your emotions. I started slowly to open up to [Jesper] to see how much he could bear to hear about my life.

After Jesper had opened up to him, Bilal also slowly began to open up to Jesper. His relationships with Jesper and the other students were growing, but around this time, Bilal was called on guard in the gang environment. Bilal is recognized by Jesper, but at the same time twisted and torn by the contradictions of meanings between his very different communities:

> I have paranoia. I've got a network at the school; I am beginning to believe in myself. Jesper is finding something in me, that I did not know I had; 'you can do this, it is legal,' I see him as a role model, as a good friend who can share his experiences of how he had felt. And sometimes he is spot on. We meet outside of school and eat together.

Bilal was interpellated into the People's College community and collectives, which include friendship, mutuality and legal and meaningful activities, but the gang community was also interpellating Bilal into the position of an armed guard. However, the practice ideology of the gang community had lost its power, it was no longer as appealing to Bilal; a contradiction has been built up, the sense of lack of meaning is getting stronger:

> Someone sticks a [weapon] in my hand, I have to stand guard for one hour. I have not seen them for a very long time. I did not feel that their problems were my problems anymore. They had inspired us to this path, but they had not told us about the ugly side. At People's College I saw what a life a normal person might have and I had the relationship to Jesper. That hour was worse than a year of isolation. You do not know if it is a civil policeman passing. I throw [the weapon] in a child's stroller, then 5 minutes later I go back and pick it up, I walk back and forth, not knowing what to do. When I return to the school I am shaken. I don't know if I can share this [experience] with Jesper. I contact the boys and tell them that I can't do this anymore – it's over.

But closing that door is also a loss:

> You have to say no to friends you grew up with, I also have to keep relatives at distance and say good-bye to my whole life. I do not have the same interests as my friends anymore. It begins to be awkward, [but] I could tell my parents that they could be proud of me, not having to defend me.

The People's College community, the new friendship with Jesper and the potential proudness of his parents wins: It is over. Bilal had moved into a boundary position, where there was too much at stake; he had too much to lose.

Recognition: Interpellation and Constitution of Bilal as Unique, Legitimate and Productive

> Recognition establishes a curious kind of suspense, since it wields the power to define the other as subject in terms whose meaning is later to appear, and on behalf of a collective that is only emerging. (Nissen, 2012, p. 170-171).

About a month later, Jesper asked Bilal to present his story along with Ali and Jesper, for the students, and he accepted. Presentation as part of a collective helped bring about an atmosphere of openness and understanding. Ali and Jesper, who also represented former criminals, had changed their lives and become unique and recognized persons, making an important difference for others like Bilal. Presenting as part of this new collective opened up for meaning-making processes and the potentiality of recognition. Bilal is not just interpellated into People's College and a friendship with Jesper, he is also slowly becoming part of a new brotherhood, representing another (masculine) practice ideology, of young adult men with criminal pasts that they are reinventing as part of new, meaningful activities, developing their unique talents, and building competence in social representation. Bilal continues to become more a part of this collective, becoming co-organizer of a course about gang criminality along with Jesper, teaching together and meeting new role models such as Robert:

> Robert, he is such a man, he truly inspires me. [...] I have some kind of engagement or competition with people. It is a bit like that with you too, [Jesper]. You can teach, and for some reason I would like to compete with you. I should be able to do better. [...] That's how I feel, also with Robert. I think he makes some damn cool stuff: movies about life in prison and gang conflict.

As part of this collective, Bilal gets to understand himself as unique and universal; like Robert, Jesper and Ali, he finds a new meaningful position to fill, a position which is not yet taken: "I'm young, and this field also needs some younger forces. Let me tell about the present, about how it is nowadays."

Already on their first meeting, when Jesper talked about the courses at the school, Bilal responded with, "you teach a course about gang criminality? I know a lot about that, I could help you teach." At that time Jesper hesitated, as it is a big thing to speak in front of so many people. Six months later, Bilal actually became the co-teacher, bringing in a new idea, as he stated, "not just presentations, power points and talking." Bilal changed the course in direction of more bodily experiences of learning: showing 'local eye' videos from gang fights outside court rooms, starting class by shouting loudly, like some gangs in Denmark, "ARE YOU FUCKING READY" – with "us" students shouting the response, "ALWAYS READY." He introduced role playing, with Bilal and Jesper being coaches of the "two gangs" in class. It was a great success, with the highest enrollment numbers of any course at the school.

Together, Bilal and Jesper are walking and expanding the margins, producing new legitimate boundary positions and new societal criteria of success (Mørck 2010), where former gang members can get recognition for

their courage, their special competences and talents through representing their story, and expanding knowledge of how to produce gang exit.

Conclusion

The goal of this paper is empirically to explore learning and life processes, when young men change their life and move beyond gang cultures, and to develop an analytical framework relevant for this purpose.

The paper applies a social practice theoretical analytical framework of expansive learning and develops it further in relation to analyze individual, collective and societal movements beyond 'gang exit'. The framework is interdisciplinary, combining social practice theory of situated learning and the historical productions of persons (such as Lave & Wenger, 1991, Holland & Lave, 2009,) with critical social psychology regarding subjects' life conduct and life trajectories (Dreier, 2008; Holzkamp, 2013), conceptualizations of collectivities as social work interventions, and the active participation and production of meaning-making processes, interpellation and belonging as part of collective subjectivities (Nissen, 2012, Mørck, 2010; 2011).

The empirical analysis of Bilal's movements challenges established notions of motivation and exit. Along with Jesper and I, he is expanding knowledge about gang exit: Movement beyond gangs is about access to new meanings and belongings. This paper hereby challenges established ideologies of motivation, especially the dualistic question of whether a candidate is "motivated" or not to exit is testable in a "motivational talk." Bilal's expansive learning is a long process which starts long before he experiences a "breaking or turning point," and it involves collectives undergoing change, not just the individual "exit candidate." Expansive learning involves questions of access to meaningful activities, belonging in communities, and things to do that oneself, one's parents, mentors and new brotherhoods can be proud of. Expansive learning is about paving new trajectories where atmospheres of openness and understanding are produced, expanded rooms for complex social representations where it is legitimate to be in process, not to be judged as either motivated or not, either gang member or not, either bad or good. This paper illustrates how criminal pasts, in these kinds of boundary communities, can be developed to become productive and open new doors of recognition into more overlapping communities. Bilal's stay at the People's College and the mentor activity were supported by his municipality, both financially and through dialogue and the belief that they were on the right track. But at the same time, his movement is also part of an enduring and collective struggle for legitimacy, because his process and the constellation of some of his new communities are in conflict with the institutionalized ideologies and normativities of good and bad in established discourse. This expansive learning of movements beyond gang life was in contradiction to 'unwritten' rules, by recognizing that the movement beyond is a process, with periods of being both a gang member and moving beyond gang life. It is a gradually change with a telos of less gang culture, and gradual changes in social self-understanding. This was also a break with the established practice ideologies, by using a mentor and role models with criminal pasts, and building up collectives that, in addition to professionals and People's College students, also include other former criminals. In other words, communities where (former and potential) gang members are invited in as legitimate participants.

Summing up, the analysis of Bilal's moments and movements beyond gang life highlights three dimensions of relevance for social practice theory, research methodology and for the development of (alternatives to) state institutionalized gang exit in Denmark:

1. It is important to understand and conceptualize the young men's (lack of) meanings and/or (lack of) belongings in and across a plurality of communities, including how the gang environment is competing with other communities. This movement beyond can be possible in boundary communities produced through new cultural activities in collaborations between people with criminal pasts, the People's Colleges and NGOs.
2. To understand and support movements beyond gang life, it is important to work more process oriented, promoting open-ended curiosity, including the understanding of embodied affects and sensations in social practice theory, methodology and intervention practice. In other words we need to move 'gang exit' practice and ideology beyond rationalist tendencies of focusing on motivation as verbalized action reasons.
3. The processes of social representation and change in social self-understandings are complex, filled with conflicts, contradictions and at the same time very important to transcending marginalization, moving all involved parties and the interventions beyond the risk of reproduction of gang labeling and demonization.

Endnotes

(1) "People's College" is a translation of the Danish "Højskole", an independent, alternative boarding school for adults.
(2) "Wild" is a term attributed to a particular segment of urban youth in Danish society - both Danish and ethnic minority youth who are involved in different "alternative," and at times illicit, activities, some becoming so called wild social street worker, coming from within the wild community and serve that community (see Mørck, 2010; 2011; 2000).
(3) The Gang Seminars are annual, dialogical seminars at Grundtvigs Højskole started by Ali and Jesper and held since 2009 (see Mørck et al, 2013).

References

Brotherton, D. C. & Barrios, L. (2004). *The Almighty Latin King and Queen Nation: Street Politics and the transformation of a New York city gang*. New York: Columbia University Press.

Davies, B. & Gannon, S. (Eds.). (2006). *Doing Collective Biography*. New York, NY: Open University Press.

Deuchar, R. (2009). *Gangs, Marginalised Youth and Social Capital*. Stoke on Trent: Trentham Books.

Dreier, O. (2008). *Psychotherapy in Everyday Life*. Cambridge, NY: Cambridge University Press.

Fine, M. (1994). Working the Hyphens. Reinventing Self and Other in Qualitative Research. In N.K. Denzin & Y.S. Lincoln (eds.), *Handbook of qualitative research* (pp. 130-154). Thousand Oaks, CA: Sage Publications.

Holland, D. & Lave, J. (2009). Social Practice Theory and the Historical Productions of Persons. *Actio: An international journal of Human Activity Theory 2009*(2), 1-15.

Holzkamp, K. (2013a). The Fiction of Learning as Administratively Plannable. In Schraube & Osterkamp (Eds.), *Psychology from the standpoint of the Subject. Selected Writings of Klaus Holzkamp* (pp. 115-132). Basingstoke : Palgrave Macmillan.

Holzkamp, K. (2013b). Psychology: Social Self-understanding on the Reasons for Action in the Conduct of Everyday Life. In Schraube & Osterkamp (Eds.), *Psychology from the standpoint of the Subject. Selected Writings of Klaus Holzkamp* (pp. 233-341). Basingstoke : Palgrave Macmillan.

Jacobsen, M.H. & Kristiansen, S. (2001). *Farligt feltarbejde – etik og etnografi i sociologien*. Århus Universitetsforlag.

Jefferson, A. M. (2004). *Confronted by practice. Towards a critical psychology of prison practices in Nigeria.* (Unpublished doctoral dissertation). Department of Psychology, University of Copenhagen.

Justitsministeriet (2011). *Rocker- og bandemedlemmer får skræddersyet vej ud af kriminalitet* [Press release]. Retrieved from http://www.justitsministeriet.dk/nyt-og-presse/pressemeddelelser/2011/rocker-og-bandemedlemmer-f%C3%A5r-skr%C3%A6ddersyet-vej-ud-af

Khawaja, I., & Mørck, L. L. (2009). Researcher positioning - when researching suspicion and Muslim otherness. *Qualitative Research in Psychology 6*(1-2), 28-45.

Kristensen, K. & Mørck, L. L. (in press, 2014). ADHD Medication and social self-understanding – Social practice research with a first grade in Danish Primary School. *European Journal of Psychology of Education.*

Lave, J. (1997). Learning, apprenticeship, social practice. *Journal of Nordic Educational Research, 17*(3), 140–151.

Lave, J. (2011): *Apprenticeship in Critical Ethnographic Practice.* Chicago: University of Chicago Press.

Lave, J. & Wenger, E. (1991). *Situated Learning. Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Mørck, L.L. (2006). *Grænsefællesskaber. Læring og overskridelse af marginalisering.* Roskilde Universitetsforlag.

Mørck, L. L. & Huniche, L. (2006) *Critical Psychology in a Danish Context. Annual Review of Critical Psychology*, 5. Retrieved from www.discourseunit.com/arcp/5.

Mørck, L. L. (2007). Young ethnic minorities in education: How to expand possibilities for learning and transcending marginalisation. In V. van Deventer (Ed.) *Citizen City. Between Constructing Agents and Constructed Agency* (149-159).

Mørck, L. L. (2010). Expansive Learning as production of community. In W. R. Penuel & K. O'Connor (Eds.), *Learning Research as a Human Science National Society for the Study of Education Yearbook, 109 (1)*, 176-191. Teachers College Press/Records.

Mørck, L. L. (2011). Studying empowerment in a socially and ethnically diverse social work community in Copenhagen, Denmark. *Ethos*, *39*(1),115-137.

Mørck, L. L., & Nissen, M. (2005). Praksisforskning. Deltagende kritik mellem mikrofonholderi og akademisk bedreviden. In T. B. Jensen & G. Christensen (Eds.), *Psykologiske & pædagogiske metoder. Kvalitative og kvantitative forskningsmetoder i praksis, Vol. 1* (pp. 123-153). Frederiksberg: Roskilde Universitetsforlag.

Mørck, L. L.; Hussain, K.; Møller-Andersen, C.; Özüpek, T.; Palm, A. & Vorbeck, I. (2013). Praxis development in relation to gang conflicts, Copenhagen, Denmark. Outlines. Critical Practice Studies 14(2), 79-105. URL: http://ojs.statsbiblioteket.dk/index.php/outlines/article/view/9794

Nissen, M. (2000). Practice Research – critical psychology in and through practices. *Annual Review of critical psychology, Action Research, 2, 145-179 (ISSN: 1464-0538).*

Nissen, M. (2012). *The subjectivity of participation – Articulating Social Work Practice with Youth in Copenhagen*. Great Britain: Palgrave Macmillan.

Osterkamp, U. (2001): Lebensführung als Problematik der Subjektwissenschaft. *Forum Kritische Psychologie 43*, 4-35.

Wenger, E. (1998). *Communities of Practice*. Cambridge: Cambridge University Press.

Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science

Niral Shah, University of California, Berkeley, Berkeley CA 94720 USA, niral@berkeley.edu
Colleen Lewis, Harvey Mudd College, Claremont, CA 91711 USA, lewis@cs.hmc.edu
Roxane Caires, New York University, New York, NY 10012 USA, roxane.caires@gmail.com

Abstract: This paper presents a comparative case study of the different ways that equity and inequity emerged as an elementary computer science student collaborated with two different classmates on programming tasks. Data collected include audio recordings of students' interactions, field notes, written assessments, and students' digital work. Using a mixed methods approach, quantitative patterns were identified in the distribution and content of student talk at multiple grain-sizes, which were analyzed in conjunction with pivotal sequences of interaction. Findings indicate that despite the existence of participation structures designed to foster equitable collaboration, inequities emerged in both dyads as students positioned themselves and their classmates with identities as more or less competent in computer science. While in the first dyad this positioning was often overt, in the second dyad positioning assumed a more passive form. Further, there is evidence that these positionings had an impact on students' opportunities to learn.

Introduction

Collaborative learning is a complex pedagogical undertaking that has the potential to foster both individual content learning and mutual respect between peers (O'Donnell & Hmelo-Silver, 2013; Webb & Farivar, 1994). However, when not carefully structured, collaborative learning can also inadvertently lead to inequitable learning situations where all students do not have access to the resources needed for learning (Cohen & Lotan, 1995; Esmonde, 2009; Sfard & Kieran, 2001). To better understand the interaction between collaborative learning and equity, this paper presents a comparative case study of how equity (and inequity) emerged as two dyads of elementary computer science students collaborated on joint programming tasks.

Learning environments are equitable when all students have access to the cognitive and social resources that foster learning given their particular histories and needs (Abrahamson & Wilensky, 2005; Boaler, 2008). From a sociocultural perspective, one resource that is critical to engagement in the learning process is a positive domain identity as a competent learner (Nasir & Hand, 2008). That is, if students are not afforded opportunities to develop such identities, or if they are actively positioned as having low intellectual status, then they may not participate in the kinds of practices (e.g., asking questions, sharing their ideas) that have been shown to facilitate content learning (Cohen & Lotan, 1995). Characterizing the equity dynamics of a collaborative learning situation, then, call for an analysis of the processes by which students position themselves and their peers as being certain types of learners, with varying potential to contribute to joint problem solving.

The findings presented in this paper are based on a mixed methods analysis of field notes, transcripts of audio data, written assessments, and digital student work collected during a three-week introductory computer science course for 10-11 year-old students in Northern California. The study focused on a student named Jason (pseudonym) as he collaborated with two classmates, Aaron and Samantha (pseudonyms), on separate days. Three strands of analysis were coordinated to progressively support or refute hypotheses related to the equitable nature of each dyad's interactions: 1) quantitative patterns were identified in how talk was *distributed* within each dyad across different phases of problem solving; 2) quantitative patterns were identified in the *content* of students' talk; and 3) qualitative analysis was conducted of key sequences of interaction in both dyads, which shed light on how students *positioned* each other (and themselves) with respect to identities as competent doers of computer science.

Findings indicate that inequities surfaced in both dyads' interactions through two different types of positioning: overt and passive. In the Jason-Aaron dyad, Aaron dominated the interaction by overtly positioning Jason as not capable of contributing to the problem solving process on multiple occasions. There is evidence that this inequity had a material impact on Jason's opportunities to learn. In contrast, data suggest that the Jason-Samantha dyad was more equitable. However, while overt forms of positioning were less prevalent in their interaction, analysis revealed instances where Samantha passively positioned Jason as less competent (e.g., by not asking for Jason's input on something she did not understand before asking a teacher for help).

These findings are significant because they illuminate aspects of the complex interaction between positioning, status, and content learning. Further, they demonstrate how even when collaborations seem equitable on the surface, students may be positioning each other in subtle ways that lead to inequitable conditions. Understanding some of the ways in which students come to be positioned as competent is critically important to designing learning environments where all students have access to identities as capable learners.

Equity, Identity, and Positioning in Social Interaction

In educational contexts, "equity" has primarily been conceptualized in terms of either institutional barriers or performance gaps on standardized tests. On the one hand, either a lack of material resources (e.g., computers, exemplary teachers) or a lack of access to the most rigorous coursework available (i.e., "tracking") can limit students' opportunities to learn (Margolis, et al., 2008; Oakes, 2005). On the other hand, even in the absence of such institutional barriers, inequities can still emerge at the classroom level as students engage with each other in the learning process.

One dimension of equity that is linked to learning concerns is students' access to the resources needed to construct identities as capable learners in a particular domain (Langer-Osuna, 2011; Nasir & Hand, 2008; Wortham, 2006). But what is meant by "identity"? According to Davie & Harré (1990), "an individual emerges through the processes of social interaction, not as a relatively fixed end product but as one who is constituted and reconstituted through the various discursive practices in which they participate" (p. 46). Thus, rather than a stable trait, "identity" can be thought of as the effect of an ongoing process of positioning embedded in social interaction (Davies & Harré, 1990; Hall, 1996).

Applying this conceptualization of identity to the context of learning puts issues of intellectual status firmly in view. That is, learning is not a neutral process. The ways in which students engage in the social practices of learning are related to the positions they occupy within the learning environment. Students that occupy positions of high intellectual status (i.e., those that signify domain competence) may take up (or be granted) opportunities to participate in the learning process more often than students that occupy positions of lower intellectual status (Cohen & Lotan, 1995). Conversely, the more students engage in the learning process (e.g., by offering ideas, answering questions), the more readily they come to occupy positions of high intellectual status. A key point is that students are not *inherently* "low" or "high" status. Rather, the types of learners that students become reflect what the learning environment makes available to them, how they are treated by their classmates and teachers, and how they understand their own capabilities.

Issues of identity formation and positioning are particularly salient in collaborative learning contexts because students are simultaneously confronted with both the cognitive challenges and social dynamics of joint problem solving (Esmonde, 2009). For this reason, the present study sought to investigate some of the ways students come to be positioned as competent (and less competent) learners in computer science while collaborating with a partner on a common set of programming tasks.

Methods

Data were collected in a university-sponsored, summer-enrichment program for academically high-achieving students entering the sixth grade. The course was titled "Creating Music, Movies, and Games with Computers" and taught introductory computer science concepts primarily using Scratch, a visual, drag-and-drop programming language. There were two sections of the course offered, which each met for 36 hours over 12 days during the same three-week period. Details about the research context and pedagogy are described in a previous paper (Shah, et al., 2013). Forty-five students were enrolled across the two sections, of which twenty-three students (51%) identified as female in enrollment records. After the first day of class, 12 students (6 per class) were selected as focal students in an effort to maximize the variation between focal students with respect to gender, race, and personality.

The course was designed and taught by the first two authors, who were supported by two teaching assistants. All instructional time was observed by at least one of three researchers who took fieldnotes focused on one or more focal students. During each 180-minute instructional day, each focal student was observed for at least 45 minutes and audio recorded for at least 90 minutes. A typical class included a 15-minute paper-based assessment (referred to here as "warm-ups"), lecture and whole-class discussion, and students working on programming tasks presented in an online curriculum. The majority of each class involved students working individually or in pairs on the computers. On odd-numbered days students worked individually, while on even-numbered days they worked in pairs on the same computer. When working in pairs, students shared a computer using a pedagogical structure common in computer science education called "pair programming" (Braught, Wahls & Marlin Eby, 2011). Pair programming utilizes roles to promote equitable collaboration (cf. Palinscar & Brown, 1984). Pair programming involves one student playing the role of "driver" while the other student plays the role of "navigator." Drivers use the keyboard and mouse, while navigators support the problem solving verbally without touching the keyboard and mouse. Further, students switched roles every five minutes so that each student has an equal opportunity to take up both roles.

Comparative Case Study Design

The full collection of 98 fieldnotes for the 12 focal students were systematically reviewed, discussed, and summarized by the research team. The summaries included perceptions of the focal students' trajectories in the class, interactions with partners and other students, and interactions with teachers. From a review of the fieldnotes and teaching experiences within the class, Jason was identified as a particularly interesting case

because of the variation we observed in his pair programming relationships. That is, because at times he appeared to be engaged in a productive collaboration, and at other times he was not, it was thought that comparing Jason's learning experience across these different contexts might put into relief features of collaboration that might help explain this variation.

Audio recordings from two days when Jason was pair programming were selected because they included acceptable audio quality and variation in the quality of his partner interactions. On Day 4 of the twelve-day class, Jason pair programmed with Aaron. Jason's average scores on homework and warm-ups were some of the lowest in the class, while Aaron had the second-highest average performance on warm-ups among the 45 participants. The content students worked on during Day 4 included programming their character to make various geometric shapes (e.g., a square "spiral" that progressively increases in size; a four-leaf flower). These tasks built on basic knowledge students learned on Day 2 about programming basic shapes, such as a triangle and a circle.

On Day 6 Jason pair programmed with Samantha, who like Aaron was also a high performer in the class: she had the eighth highest average performance on warm-ups among the 45 participants. The content worked on during Day 6 involved making a game of "tag," which involved students learning to use conditional statements (e.g., "if" blocks).

Transcripts were created of all audible dialogue in the audio recordings. Each transcript was divided into turns, which were differentiated by either a new sentence or a distinctly new idea. Transcripts included few indications of tone, but documented phrases interpreted as questions using a question mark.

Analytical Approach

The present study characterized an equitable collaboration as one where all group members have an equal opportunity to participate in the interaction, as well as to be positioned as having high intellectual status (cf. Cohen & Lotan, 1995). Prior research has used student talk as a measure of engagement in collaborative settings (see Barron, 2003; Engle & Conant, 2002; Langer-Osuna, 2011). As some have argued, researchers should be careful not to over-privilege talk, particularly in light of research showing that nonverbal behavior (e.g., eye gaze) is also a meaningful indicator of engagement (Sawyer, 2013). However, the lack of high-quality video constrained the possibility of such an analysis. Still, much can be learned about the equity in an interaction through quantitative and qualitative analysis of verbal communication. Overall, our analytical approach was consistent with Barron's (2003) recommendation to emphasize the group itself as the unit of analysis.

Analysis using mixed methods was conducted at three levels. The first level focused on the *distribution* of talk within dyads in three areas of interaction: total overall talk, across the driver/navigator participation structure, and across certain classroom activities. In general, it was assumed that if the collaboration were equitable, then the number of turns would be near equally distributed (i.e., 50-50). In addition to determining the overall distribution of talk for the entire 90-minute interaction, transcripts were divided into six classroom "activities": a) whole class; b) interacting with teachers; c) coding; d) designing; e) managing logistics; and f) off-task. Each turn in the transcript was tagged with one of these six activities, and then talk-distribution was analyzed for each activity. The rationale for this part of the analysis was that only examining overall talk-distribution might obscure inequities at a finer grain-size in activities like "coding," which was particularly relevant to the problem solving process.

The exception to the 50-50 benchmark was the distribution of talk within the driver/navigator participation structure. That is, because navigators were expected to talk more given that this was the nature of the role, a 50-50 distribution of talk would not necessarily be desirable. Instead, the equity benchmark here was that the interaction would exhibit a "mirroring" effect, such that if talk distribution is 70-30 when Student A is navigating and Student B is driving, then when the students switch roles (i.e., Student A is driving and Student B is navigating) the distribution of talk should also switch (i.e., 30-70). This mirroring behavior could be interpreted as evidence that students are honoring the intent of the participation structure, which was to promote fairness in their social interaction.

The second level of analysis concerned the *content* of students' talk. Based on a review of coding schemes in the collaborative learning literature, as well as a process of open coding informed by our field notes that helped identify types of student talk that might be consequential for positioning vis-à-vis intellectual status, a coding scheme was developed that aimed to capture the quality of student talk. While the authors' cultural interpretations of students' statements informed the development of scheme, the codes were designed to describe the content of the talk rather than our perception of the likely effects. After multiple rounds of refinement on a subset of the transcripts, the codes stabilized and all of the turns were coded with the transcripts divided evenly between the first two authors. Each turn in the transcripts could be tagged with multiple codes.

While some of the codes occurred with high frequency (e.g., "asking a question," "issuing a command"), other codes occurred with lower frequency (e.g., "complementing a partner," "asking a teacher for help without consulting partner") but were still considered potentially consequential for positioning. High frequency codes were analyzed to determine how a given type of statement was distributed in a dyad. For

example, if one of the two students issued disproportionately more commands, then it might suggest an inequity in the collaboration. Due to space constraints, the findings presented in this paper focus on the two highest frequency codes in the data: "student questions" and "commands" (subsequently described in greater detail).

The third level of analysis involved a qualitative analysis of key sequences of interaction that were both representative of the interaction and that seemed consequential for positioning and learning. Qualitative approaches are particularly useful in revealing the processes by which overall interactional trends emerge (Sawyer, 2013). In this case, the focus was on pivotal moments in each dyad where students were positioned either as competent doers of computer science or as having low intellectual status.

Findings

This section presents three levels of analysis: 1) distribution of talk; 2) content of talk; and 3) fine-grained processes of positioning. Each level involves multiple strands of analysis, which were coordinated to generate and then progressively refine hypotheses about how equitable both dyads were in various facets of their interactions.

Distribution of Talk within Dyads

An initial indicator of an equitable collaboration is whether the distribution of talk across the entire 90-minute interaction within dyads approached a 50-50 distribution. Figure 1 shows that while Jason engaged in fewer turns than Aaron overall (N=991; J:41%; A:59%), Jason and Samantha engaged in nearly the same number of turns (N=1000; J:51%; S:49%). In terms of an initial hypothesis, these data suggest that the Jason-Aaron dyad may have been less equitable than the Jason-Samantha dyad.



Figure 1. Distribution of talk within dyads.

Distribution of Talk within Pair Programming Participation Structure

A second indicator of an equitable collaboration in this context is whether dyads were respecting the norms of the driver/navigator participation structure. As described earlier, this would mean that: 1) navigators were talking more often than drivers, and 2) that the distribution of talk would exhibit the "mirroring" effect (i.e., when students switch roles the distribution of talk would also switch and be roughly equivalent).

Figure 2 shows how the distribution of talk within each dyad compared when each student was acting as the navigator. Data indicate that the Jason-Samantha dyad satisfied both criteria of an equitable collaboration within this participation structure: the navigator always talks more, and the distribution of talk exhibits the "mirroring" effect (see Figure 2). When Jason was the navigator, he engaged in 56% of the turns (N=415; J:56%; S:44%); when Samantha was the navigator, she also engaged in 56% of the turns (N=353; J:44%; S:56%). This finding supports the hypothesis that the Jason-Samantha dyad was equitable.

In contrast, the Jason-Aaron dyad did not satisfy either criterion. When Jason was navigating he actually spoke slightly *less* often than Aaron (N=341; J:49%; A:51%). And when Aaron was navigating, not only did he engage in a greater proportion of turns (N=574; J:34%; A:66%), but roughly 200 more turns occurred overall. Altogether, these data suggest that Aaron was dominating the interaction, and lend further evidence to support the hypothesis that the Jason-Aaron dyad was less than equitable.

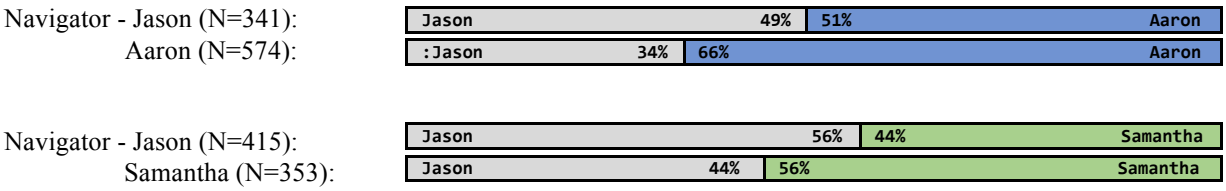


Figure 2. Distribution of talk within pair programming participation structure.

Distribution of Talk across Activity Contexts

A third indicator of an equitable collaboration is how talk was distributed within the finer grain-sizes of activity that students engaged in during the problem solving process. The analysis here focuses on two activity contexts: "coding" (e.g., generating new code, debugging) and "designing" (e.g., planning their project, making aesthetic choices about what their characters should look like). Not only were these two activities the most common during each dyad's interaction, it was assumed that "coding" and "designing" would be especially relevant for positioning and students' opportunities to learn.

Figure 3 shows the distribution of talk when each dyad was engaged in coding. In the Jason-Aaron dyad, the data show that Aaron dominated the interaction (N=492; J:34%; A:66%). Because "coding" was where much of the content of students' work was generated, the imbalance in this activity context may have been particularly impactful in positioning Aaron as high-status (i.e., the one who knows the answers) and Jason as low-status. In contrast, the distribution of talk in the Jason-Samantha dyad during coding was almost equal (N=485; J:48%; S:52%). Both data support the running hypotheses about the equity of each dyad's interactions.

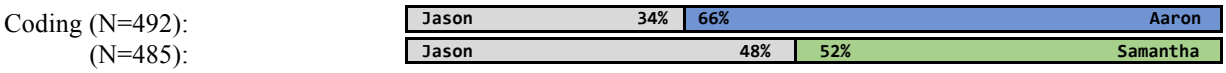


Figure 3. Distribution of talk when coding.

Figure 4 shows the distribution of talk when each dyad was engaged in project design. Interestingly, the analysis reveals that this is one context where Aaron did *not* dominate the interaction (N=328; J:48%; A:52%). This complexifies the current hypothesis about the Jason-Aaron dyad in that it shows that there was at least one part of the interaction that *did* appear equitable. In the Jason-Samantha dyad, the data show that Jason actually spoke more than Samantha during this phase of problem solving (N=163; J:55%; S:45%). However, because "designing" involved the aesthetics—as opposed to the content—of students' projects, it is likely that it was less consequential than "coding" in positioning students as academically capable or "smart." Still, it is noteworthy that Jason engaged in a greater proportion of turns during this particular phase of problem solving.

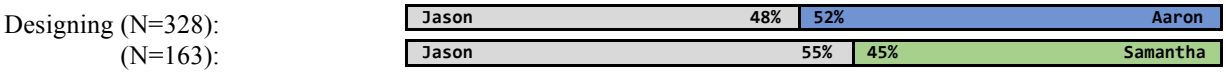


Figure 4. Distribution of talk when designing.

Content of Talk in Dyads

Complementing the previous analyses, the content of student talk was also analyzed to shed light on the nature of students' interactions. This section focuses on the two codes that occurred most frequently: "commands" and "questions." In the next section some of the less frequent—but potentially consequential—codes are discussed.

Note that while in a colloquial sense "commands" have a typically pejorative connotation, the operational definition used here was more neutral and referred to any turn in which one student directs her/his partner to execute a particular action. Still, issuing many directive statements may indicate a lack of respect for the intellectual capacity of that individual to contribute. Figure 5 shows the number of commands issued by students in each dyad. In the Jason-Aaron dyad, Aaron issued nearly 100% of the commands (N=123; J:5%; A:95%). This finding adds strong support to the hypothesis that their interaction was inequitable. While the pattern was not as pronounced in the Jason-Samantha dyad, Samantha did disproportionately issue commands at roughly a 2:1 ratio relative to Jason (N=71; J:35%; S:65%). However, the total number of commands was less overall compared with the Jason-Aaron dyad (71 vs. 123 commands). This is the first indicator that the Jason-Samantha collaboration may not have been as equitable as the analysis has indicated.

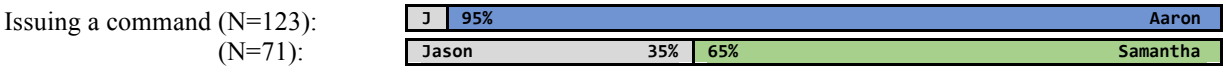


Figure 5. Distribution of commands.

In addition to commands, the act of asking questions to a partner is also relevant to positioning because it presumes that the partner can answer it. In other words, asking a partner a question can be seen as actively positioning that person as competent and, alternatively, *not* asking a partner a question can be seen as passively positioning that person as *less* competent. Figure 6 shows that in both dyads, Jason asked more questions than both Aaron (N=42; J:67%; A:33%) and Samantha (N=41; J:68%; S:32%) by almost identical 2:1 ratios. An environment where students are asking each other questions may still be equitable if question-asking is reciprocal. However, because question-asking was imbalanced in both dyads, it bolsters the hypothesis that the Jason-Aaron dyad was generally inequitable, and it further challenges the hypothesis that Jason-Samantha was an equitable collaboration.

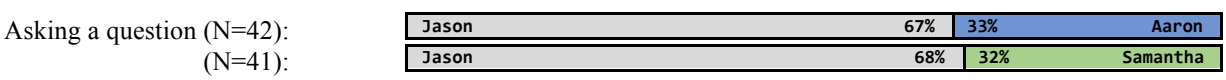


Figure 6. Distribution of questions.

Processes of Positioning

Key Sequences in the Jason-Aaron Dyad

The preceding analysis showed that Jason asked more questions than Aaron, and that Aaron frequently commanded Jason. Based upon our assumptions that questions positively position the partner who is asked, and that commands negatively position the partner who is being commanded, these data suggest that the Jason-Aaron partnership was inequitable. The qualitative analysis presented here demonstrates both how Jason is positioned as not competent through the use of commands, and also highlights how Jason's interactions with Aaron were consequential for his opportunities to learn.

A typical pattern in the Jason-Aaron dyad was a "command and clarify" routine: when Aaron was navigating that he would issue commands to Jason, and Jason would intermittently ask clarifying questions. To demonstrate this pattern, consider the following sequence of turns in which Aaron provides instructions to Jason to find a "When Space Key Pressed" block from the "Control" tab in the Scratch user interface, and to change the "space" key to the letter "D" key.

168 Jason: So what next?
169 Aaron: When "D" key pressed.
170 Jason: "D"…what?
171 Aaron: Control tab.
172 Jason: Oh…here?
173 Jason: Just press "D"?
174 Aaron: Don't just press "D," go to the Control tab.
175 Aaron: "When Space Key Pressed" [block]…do you see that?
176 Aaron: Space key pressed.
177 Jason: This?
178 Aaron: Yeah, drag it out.
179 Aaron: Change it to "D."

With Aaron issuing repeated directives and Jason responding with clarifying questions, Aaron is positioned as the "competent" partner and Jason is positioned as having low intellectual status. This inequitable dynamic relates to Jason's opportunities to learn. For example, there was evidence that Jason was not always engaging with the content of Aaron's commands, even when Aaron's commands would not produce the correct solution. Following the exchange detailed above, the students are attempting to write a program that will produce a 5-pointed star. To produce this geometric shape, the sprite must turn 144 degrees at each vertex. Students had not encountered the problem before, so it was not surprising when Aaron initially suggested that the turning-angle be 108 degrees:

190 Aaron: Turn one hundred eight degrees.
191 Jason: Turn – one hundred and. [speaking aloud while typing on the keyboard]
192 Jason: Eight.
193 Aaron: That's one-eighty [pointing out that Jason typed in 180, not 108]
194 Jason: Oh wait – you want one hundred and eight?
195 Aaron: Yeah.

In this interaction, Aaron instructed Jason to use the number 108, but Jason instead types in 180. Although Aaron's suggestion of 108 degrees will not successfully draw the shape, it is a reasonable estimate of how much the sprite should turn. In contract, turning 180 degrees is not reasonable because this would cause the sprite to face the opposite direction. Given students' prior experiences with Scratch-based geometry on previous activities that day and two days earlier on Day 2, we would expect that students could easily determine that 180 degrees should not be used in creating the shape. Jason's statement in line 194 ("Oh wait – you want one hundred and eight?") suggests that either he did not perceive 180 as an unreasonable solution or did not feel he had the authority to challenge Aaron. As a result, Aaron's incorrect estimate was left unchallenged, thereby impeding both students' problem solving.

There is further evidence that this type of uncritical execution of Aaron's commands hindered Jason's opportunities to learn, in particular. On the following day of class, Jason requested teacher help for a review problem that Aaron and he had completed the previous day. Jason reported to the teacher that he did not remember how to solve the problem *"I don't remember how to do it, [because] when we did—well I didn't do it, Aaron did. I was there but I didn't really like..."* (lines 287-288). Unfortunately, audio recordings are unavailable from when they completed the task for the first time. Jason's statement that he does not know how

to solve the problem may be accurate or reflect his lack of confidence; however, both could plausibly have resulted from the negative positioning he experienced during his collaborative work with Aaron.

Key Sequences in the Jason-Samantha Dyad

While considering only the distribution of talk, the Jason-Samantha dyad appeared equitable. However, analyzing the content of talk revealed an imbalance between them, as Samantha issued more commands and Jason asked more questions. These two patterns may have positioned Jason as less competent. The preceding section discussed how Aaron overtly positioned Jason as not competent by directing his actions with commands. In the Jason-Samantha dyad, however, there were moments where positioning was more passive.

For example, recall that on Day 6 students were programming a game of "tag," where one sprite would earn points if it touched the other the sprite. On one occasion the curriculum asked students to write code for a "power-up," which is a special sprite in the game of tag that would afford a player bonus powers if touched. When Jason and Samantha reached this part of the curriculum, Samantha immediately calls a teacher over for help without consulting her partner. Samantha did not understand the concept of a "power-up." As the teacher explains the power-up, Jason repeatedly interjects comments explaining the idea of a "power-up." It appears that Jason could have been a resource to Samantha if she had asked him. After the exchange with the teacher, Samantha asks Jason: "You get it?" (line 127). This further indicates that she was not attending to Jason's contributions, in that she did not realize that Jason understood the concept of a "power-up" all along. However, this question may still show some level of investment in her partner's learning.

There were also three occasions (lines 347, 366, 574) where Jason was driving but Samantha made a bid to use the keyboard or mouse. In these moments, Samantha seemed impatient with how quickly Jason was typing or perusing the Scratch interface. Recall that one of the purposes of the driver/navigator participation structure was to fairly distribute access to the computer, which was highly valued by all students in the class. Only students in the driver role were supposed to use the keyboard and mouse. Thus, Samantha's bid for the computer (i.e., to "drive") can be interpreted as an implicit refutation of Jason's ability to handle the computer. So although such moments in their interaction were infrequent, they may have served to position Jason as a less than equal contributor.

Discussion and Implications

This paper offered an analysis of how equity and inequity can emerge on multiple levels as students interact in collaborative learning situations. In the Jason-Aaron dyad, Aaron dominated the conversational floor and overtly positioned Jason as less capable of contributing to their joint problem solving. In contrast, the Jason-Samantha dyad appeared more equitable (e.g., they tended to share the conversational floor). However, analysis revealed subtle inequities in some of the more passive ways in which Samantha also positioned Jason as less competent.

The Jason-Samantha case, in particular, suggests that rather than conceptualizing equity as a binary phenomenon (i.e., an interaction is deemed equitable or inequitable), it may be more useful to conceptualize equity as contextual. That is, in any given interaction there may be situations where a student had more or less access to the resources needed for learning. In fact, recall that even within the Jason-Aaron dyad, the distribution of talk was roughly equal during the "designing" activity context. Although "designing" may not be as high-status as "coding," it is possible that the opportunity for Jason to engage within this activity may have prevented him from disengaging altogether from an interaction where his partner was dominating. A different computer science classroom in which students had no opportunities for "designing" may have provided even fewer opportunities for Jason to participate in his collaboration with Aaron.

A limitation of this research is that it took place in the privileged context of an opt-in summer program for high-performing students. But while the students in this study may not be representative, it is plausible that the mechanisms of positioning identified in this paper would also occur in other educational contexts. Another consideration is that the analysis of positioning and equity presented here was confined mainly to the local learning environment. However, social interactions between learners are also situated within broader societal discourses (e.g., gender, race) that also can affect how students are positioned, especially in technically oriented domains where stereotypes about the capabilities of different groups exist (Shah, 2012; Langer-Osuna, 2011). Future research in this area should aim to account for how such discourses mediate student interaction.

There is growing interest in the learning sciences in issues of equity. To date, though, the field has few tools for measuring equity in a collaborative learning context. And indeed, there is no consensus standard by which researchers can definitively determine the degree to which a given learning situation was equitable. Still, the methodological approach presented here aims to contribute one way of operationalizing equity through quantitative and qualitative measures. Efforts to specify what we mean by "equity" can only serve to facilitate progress toward the long-term goal of creating learning environments that foster student agency and preserve all students' opportunities to learn.

References

Abrahamson, D. & Wilensky, U. (2005). Collaboration and equity in classroom activities using Statistics as Multi-participant Learning Environment Resource (S.A.M.P.L.E.R). Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*(3), 307-359.

Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal*, *34*(2), 167-194.

Braught, G., Wahls, T., & Marlin Eby, L. (2011). The case for pair programming in the computer science classroom. *ACM Transactions on Computing Education,* 11, 1–21.

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal*, 32(1), 99-120.

Davies, B., & Harré, R. (1990). Positioning: The discursive production of selves. *Journal for the Theory of Social Behavior*, *20*, 43–63.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, *20*(4), 399-483.

Esmonde, I. (2009). Mathematics learning in groups: Analyzing equity in two cooperative activity structures. *Journal of the Learning Sciences*, *18*(2), 247-284.

Hall, S. (1996). Who needs identity? In S. Hall & P. du Gay (Eds.), *Questions of cultural identity* (pp. 1-17). London: Sage Publications.

Langer-Osuna, J. M. (2011). How Brianna became bossy and Kofi came out smart: Understanding the trajectories of identity and engagement for two group leaders in a project-based mathematics classroom. *Canadian Journal of Science, Mathematics and Technology Education*, *11*(3), 207-225.

Margolis, M., Estrella, R., Goode, J., Jellison Holme, J., & Nao, K. (2008). Stuck in the shallow end: Education, race, and computing. Cambridge, MA: MIT Press.

Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences, 17*(2), 143-179.

O'Donnell, A.. & Hmelo-Silver, C. E. (2013). What is collaborative learning? An overview. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.). *The international handbook of collaborative learning* (pp. 1-15). New York, NY: Routledge.

Oakes, J. (2005). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Sawyer, R. K. (2013). Qualitative methodologies for studying small groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.). *The international handbook of collaborative learning* (pp. 1-15). New York, NY: Routledge.

Sfard, A., & Kieran, C. (2001). Cognition as communication: Rethinking learning-by-talking through multi-faceted analysis of students' mathematical interactions. *Mind, Culture, and Activity*, *8*(1), 42-76.

Shah, N., Lewis, C. M., Caires, R., Khan, N., Qureshi, A., Ehsanipour, D., & Gupta, N. (2013). Building equitable computer science classrooms: Elements of a teaching approach. *ACM SIGCSE Bulletin, 44*(1), 263-268.

Shah, N. (2012). Mathematics learning in a racial context: Unpacking students' reasoning about "Asians are good at math." In J. van Aalst, K. Thompson, M. J. Jacobson & P. Reimann (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012)* (Vol. 2, pp. 222-226). Sydney, NSW, Australia: International Society of the Learning Sciences.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal*, *31*(2), 369-395.

Wortham, S. (2006). Learning identity: The joint emergence of social identification and academic learning. New York, NY: Cambridge University Press.

Hands-on Small Group versus Whole Class Use of an Interactive Simulation: Qualitative Comparisons

A. Lynn Stephens and John J. Clement, University of Massachusetts-Amherst, Amherst, MA 01003
Email: lstephens@educ.umass.edu, clement@educ.umass.edu

Abstract: Assumptions about the superiority of hands-on use of computer simulations over projecting them in whole class have seldom been tested. Contrary to expectations, preliminary pre-post results from two lesson sequences yielded no evidence for an advantage for students in the hands-on condition. We conduct qualitative analyses of one of the lesson sequences, in which a popular simulation was used in eight high school physics class sections, half in whole class discussion and half in small groups. Videotape and activity sheet analyses of such factors as percentage of time spent on conceptual difficulties and amount of support for using key visual features of the simulation yielded no evidence for an advantage for small group students. No small group students in lower level physics sections showed evidence in written or drawn work for having utilized key visual features. A balance of complementary small group and whole class use is recommended.

Introduction

Studies have suggested that students benefit from control of the pace of animations; the speed of a presentation needs to match the speed of comprehension of a topic (e.g., Mayer and Chandler, 2001). Small group work has been prized for allowing such control as well as for providing students opportunities to interact with others, to create metaphors that other students can readily understand, and to enter engaged exploration of the concepts of a lesson. On the other hand, we have at times noticed scenarios such as the following: Above-average high school physics students were working at a computer with an interactive simulation. When one student raised an important conceptual issue concerning a question on the activity sheet, another student suggested that the first was over-interpreting the question (he wasn't). The forceful comment appeared to shut down a potentially fruitful discussion and the conceptual issue was never discussed. Observations such as this one raise the question of whether a teacher might have been able to encourage discussion of the student's question in whole class discussion, and more generally, whether there are complementary advantages to small and large group formats for discussing simulations. Here we report on a set of comparative case studies in the two formats.

Theoretical Background

A number of studies have investigated the effects of instructional guidance for simulations when the guidance was provided within the learning materials themselves (Perkins, et al., 2006; review by Cook, 2006) and the effectiveness of animations or simulations when teachers provided the verbal information (Russell & Kozma, 2005) or when at least part of the use of the simulations was in the context of whole class discussion (Raghavan, Sartoris, & Glaser, 1998). The use of simulations in small groups and by individual students has been studied (e.g., Adams et al., 2008; Linn, 2003; Reid, Zhang, & Chen, 2003; Williams, Linn, Ammon, & Gearhart, 2004). However, there do not appear to be many studies that investigate how best to support students when simulations are used in a whole class setting. Hands-on activity afforded by small group work would appear to offer students a more active learning experience with simulations than would a whole class format. In the context of think-aloud interviews, Adams, et al. (2008), felt that simulations can be highly effective only if the student's interaction is directed by the student's own questioning. This kind of self-directed interaction with a simulation would seem to require a lesson format with hands-on opportunities. On the other hand, studies have reported a variety of issues concerning the effective use of small group discussions in science classes, such as the fact that students can exhibit a low level of engagement with tasks (Bennett, Hogarth, Lubben, Campbell, & Robinson, 2010). Two studies that each compared a single small group class with a single whole class discussion (Wu & Huang, 2007; Smetana & Bell, 2009) did not find significant differences in pre-post gains. These studies, published after we had begun our project, suggest the importance of conducting a larger study.

Summary of Preliminary Results

In an earlier study (Stephens, 2012), a preliminary analysis was conducted of short answer pre-post results from two high school physics lesson sequences that used simulations and animations in hands-on and whole class contexts. We will briefly review these results. The pre-post tests consisted of transfer questions; these targeted the concepts of the lessons via questions that had not been directly addressed during instruction. Tests were administered immediately before and after instruction.

In Tables 1 and 2, College Prep (CP) was the least advanced physics level included in the study, Honors Physics (HP) was an intermediate level, and Advanced Placement (AP) the most advanced. WC and SG

 © ISLS

refer to lessons that used interactive simulations in Whole Class and Small Group formats, respectively. Given the nature and sizes of the samples, we did not attempt to extrapolate to a larger population or even to compare one matched set with another; however, we found the results intriguing enough to motivate the case study analyses that constitute the main study.

Table 1: Gravitational PE short answer transfer question pre-post gains.

	WC Gains			SG Gains					
	N	Mean	SD	N	Mean	SD	*t*-value	Sig.	Cohen's *d*
CP	11	**0.26**	0.20	14	0.25	0.24	0.097	0.924	0.04
HP	20	**0.22**	0.21	19	0.09	0.15	2.221	0.033**	0.71
AP	23	**0.10**	0.12	21	0.02	0.11	2.368	0.023*	0.71
AP	21	**0.09**	0.16	21	0.07	0.10	0.506	0.616	0.16

Notes: Boldface indicates the larger mean gain within each matched set. *Significant difference in gains in favor of the whole class condition. **Significant difference in gains in favor of the whole class condition; however, unanticipated events may have had a disproportionate effect on the small group condition.

Table 2: Projectile Motion short answer transfer question pre-post gains.

	WC Gains			SG Gains					
	N	Mean	SD	N	Mean	SD	*t*-value	Sig.	Cohen's *d*
CP	14	**0.31**	0.29	9	0.27	0.29	0.336	0.741	0.14
HP	21	0.35	0.35	25	**0.36**	0.34	-0.087	0.931	0.03
HP	34	**0.35**	0.29	19	0.32	0.24	0.294	0.770	0.08
HP	15	**0.41**	0.22	22	0.37	0.33	0.471	0.640	0.16
AP	20	0.22	0.22	21	0.22	0.23	0.036	0.971	0.01

Note: Boldface indicates the larger mean gain within each matched set.

To the surprise of the teachers in the study, in no comparison was there a significant advantage for the small group condition. If anything, in the Gravitational PE lesson sequence, there appeared to be a slight trend in favor of the whole class condition. Throughout the two years in which these sequences were conducted, the teachers continued to predict superiority for small group work over whole class work. After the second year, one of the teachers stated in an interview, "When at all possible, most of the time simulations are better done with students working in (small) groups." Thus, the results raised questions for us and for the teachers. The present study uses qualitative analyses of classroom videotapes from the Gravitational PE sequence to look for possible differences between small group and whole class interactions. Analyses of selected answers on student activity sheets shed further light on differences in student learning. (Qualitative analyses of the Projectile Motion lesson sequence will be discussed elsewhere; Stephens & Clement, in preparation.)

Qualitative Study Methods

We use comparative case study analysis of the Gravitational PE lessons, which had been conducted in whole class and small group formats, to shed light on the following overall question: *Why did the whole class format produce gains as strong as those of the hands-on small group format for these classes?*

Research Questions: For Both Conditions

1. To what extent did students and teachers engage in discussion about certain key concepts while working with the simulation?
2. To what extent did teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulation?
3. To what extent did teachers and students support the recognition, use, and interpretation of certain key visual features of the simulation?
4. Did students recognize and use key visual features of the simulation?
5. Cutting across the above four themes: Is there a difference in the way the above issues played out in the whole class and small group formats used in these class sections?

Data Sources and Collection

The intention in our classroom observations was not to train teachers to use these tools in a particular way, but to study how these teachers naturally used the tools in two common classroom formats. The Gravitational PE lesson sequence involved two teachers at a high school in a suburban college town. The teachers were purposefully selected; they had to be willing to teach model-based lessons, to foster discussions in both whole

class and small group settings, and to use computer simulations as part of their lesson plans. Class sections taught by each teacher were purposefully selected for analysis according to whether they fit criteria for matched sets, as follows. The teacher must have been teaching at least two comparable sections in a given semester and been willing to conduct the lesson sequence in at least one section in a whole class format and in at least one other section in a small group format. Teachers' evaluations and records were relied upon to determine that the sections within a set had students comparable in terms of age and demonstrated levels of aptitude for the content as evidenced by their prior work in the course. In addition, the classes in each section must have provided similar levels of preparedness for the lesson as indicated by the teachers' records of their lesson plans. Finally, the sequence as taught in the two formats must have been similar (see Materials and Procedure below) and the class sections must have been allowed similar amounts of time on the lessons and pre- and post-tests. Fifteen lesson sequences were observed; seven sequences and one teacher were dropped from analysis because they did not meet the above criteria, leaving eight sequences from two teachers to be subjected to comparative analysis. These comprised four matched sets of class sections, $N = 150$. Once it was determined that sections were matched, they were assigned to the whole class (WC) or small group (SG) condition according to practical logistical considerations. Class times rotated; on some days the teacher taught the whole class condition first and on other days the small group condition first. Teacher A taught this as a two-day sequence while Teacher B taught it as a one-day lesson. Each lesson was videotaped and one or both authors observed all lessons.

Materials and Procedure

Although materials varied slightly for each level of physics, for the two conditions within each matched set, the teacher used the identical simulation and other materials but varied the way in which the simulation was used. In the whole class condition, each teacher used a single computer to project the visuals onto a screen in front of the class and facilitated a whole class discussion as students worked through the activity sheet. In the small group condition, multiple computer stations were used with 2-4 students to a computer; they engaged in hands-on exploration and small group discussion supported by the activity sheet while the teacher circulated among the groups. In both conditions, the teacher introduced the computer activity to the whole class. In both conditions, the teacher was available for questions the entire time the simulation was in use. Other than the constraints provided by the technological set-up, the activity sheets, the simulation mode (whole class or small group) and the data-collection needs of the study, teachers were free to conduct their classes as they saw fit and were encouraged to use the best teaching strategies they could devise. Time-on-task was controlled within each matched set by using the same activity sheets and number of class periods. Though early versions of the materials were inspired by sample lesson plans from the PhET website, the final lesson plans and activity sheets, designed to work with both whole class and small group formats, were largely the construction of the teachers.

The teachers selected a simulation ahead of time from freely available online sources. They chose a sophisticated simulation developed by a research group, *Energy Skate Park* at http://PhET.colorado.edu (Perkins, et al., 2006). See Figure 1. The track can be added to or reshaped, the skater placed anywhere in the scene and released, and the simulation run to see how the skater would respond under the influence of gravity. Activity sheets guided students through an exploration of the skater's motion, the changes in the skater's potential, kinetic, thermal, and total energy with time, and the relationships between those changes. The sheets explicitly instructed students to turn on the Gravitational Potential Energy (GPE) Reference Line (the dotted line in Figure 1) and to move it around. It also instructed students to turn on the animated Energy Bar Graph (to the right in Figure 1), which showed clearly when the potential energy of the skater took on negative values.

Method of Videotape Analysis

Videotape analysis was used to develop a picture of what an individual hypothetical student could have been exposed to in a given lesson. In this analysis the video camera can be viewed as a proxy for an individual student; that is, it took the viewpoint of a hypothetical student in that classroom and recorded what she might

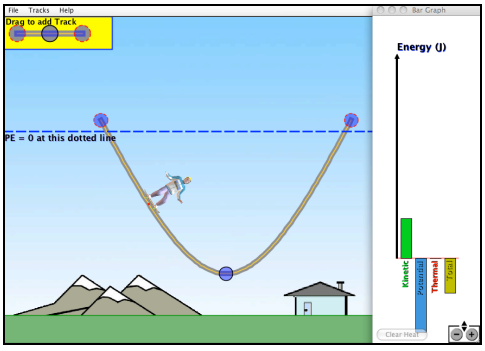


Figure 1. PhET Energy Skate Park with two key features turned on.

have seen and heard. In small group classes, at the point that the students moved into small groups, the camera moved to one of the groups also. Although fewer students were visible on camera than in the whole class condition, the videotape recorded what an individual student in that group could have seen and heard. We began analysis by using a constant comparative method to identify key behaviors observable in videotapes and transcripts of the first four classroom discussions that occurred during use of the PhET Energy Skate Park simulation. Observation categories developed from this procedure were honed in an iterative process along with coding criteria for assigning video segments to categories. This honing process constituted a major part of the effort involved in this study; criteria were developed, applied to fresh transcript sections, critiqued by a second researcher, then refined until the observation categories and their coding criteria stabilized. Finally, the criteria for the stable categories were used to code the entire simulation portion of the lesson in all eight transcripts.

Method of Analysis for Selected Questions on the Activity Sheet

A different lens is provided by activity worksheet analysis. This analysis has the strength of including work from almost all of the students in the classes and is not restricted to students who spoke on camera. However, student drawing and writing abilities varied widely and some activity sheets were difficult to interpret or were not completed. This analysis provided an estimate of how many students actually used certain visual features in their own thinking, as evidenced by their written and drawn answers to selected activity sheet questions. We began analysis by using a constant comparative method to code student responses to selected questions in a stratified sample of 30 activity sheets. Questions were selected that 1) addressed key concepts that the key visual features were thought to support; 2) asked for open-ended written and drawn answers. Coding categories developed from this procedure were honed in an iterative process with critiques from a second researcher. These were used to code student responses to the selected questions on all activity sheets for which there were legible answers (135 out of 150 sheets). Activity sheet coding was done blind to whole class or small group condition.

Results

Videotape Analysis

To What Extent Did Students and Teachers Engage in Discussion about Certain Key Concepts While Working with the Simulation?

From pilot lessons, a fundamental concept that we observed posing difficulty for the pilot students was the concept of the existence of negative energy quantities, especially negative total energy.

Code: Student or teacher mentions possibility of total energy of some system being zero.
Code: Student or teacher mentions possibility of some kind of energy value being negative.

Percentage of discussion time spent on these concepts was determined. The results provide an estimate of what an individual student in the position of the camera could have been exposed to during the discussion.

Table 3: Projectile Motion short answer transfer question pre-post gains.

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	4.32 min / 42.42 min = **0.10**	0.40 min / 23.90 min = 0.02
Yr 1 HP	Teacher A	2.85 min / 62.03 min = **0.05**	0.75 min / 29.23 min = 0.03
Yr 1 AP	Teacher B	0.92 min / 41.10 min = 0.02	0.99 min / 32.32 min = **0.03**
Yr 2 AP	Teacher B	2.58 min / 41.71 min = **0.06**	1.16 min / 28.95 min = 0.04

Notes: Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement

The percentage of discussion time spent on these concepts is shown in Table 3. Notably:

- Discussion about these two concepts ranged from 2% to 4% of discussion time in the small groups on camera and from 2% to 10% in the whole class discussions.
- Small groups spent less time on the lesson, not because less time was allowed but because they chose to finish early, thinking they were done with the activity. Therefore, the *total* amount of time spent on these key concepts was substantially less in the small group discussions than in the matched whole class discussions, ranging from less than half a minute to a little over a minute in the small group discussions and from a minute to over four minutes in the whole class discussions.

In each condition there was little discussion time devoted to these concepts. This was surprising, given that the animation provided important potential affordances for developing the concepts, including the two features shown in Figure 1, and the fact that students occasionally expressed frustration concerning these ideas. These

are not the only important concepts necessary for students to understand the material; however, they were of particular interest because their lack appeared to constitute a block to acquiring other concepts of the lesson. The evidence described here does not suggest an advantage for the students in the small group condition regarding a chance to address these stumbling blocks. Even if the quality of discussion had been much higher in the small groups than in the whole class discussions, it is doubtful that less than half a minute of discussion, as in the lower level small group, would have been sufficient to explore the concept of zero or negative energy.

To What Extent Did Teachers and Students Attempt to Respond to Conceptual Difficulties and Misconceptions During Work with the Simulation?

Students sometimes expressed frustration, confusion, or puzzlement in connection with ideas presented within the animation, the activity sheet, or the class discussion, including (but not limited to) the key concepts described above. At other times, students appeared to try to address each other's misconceptions.

Code: Response to conceptual difficulty: Classroom activity following a student expression of conceptual difficulty was considered a response if it bore some relationship to the expressed difficulty.

Code: Response to misconception: Classroom activity was considered a response to a misconception if it appeared to be an attempt by teacher or student to address a misconception.

Total time spent on such discussion was noted. The results provide an estimate of what an individual student in the position of the camera could have been exposed to during the discussion.

Table 4: Response to conceptual difficulties (expressed as percentage of discussion time).

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	6.15 min / 42.42 min = **0.14**	0.89 min / 23.90 min = 0.04
Yr 1 HP	Teacher A	14.05 min / 62.03 min = **0.23**	3.35 min / 29.23 min = 0.11
Yr 1 AP	Teacher B	3.72 min / 41.10 min = **0.09**	1.58 min / 32.32 min = 0.05
Yr 2 AP	Teacher B	1.79 min / 41.71 min = 0.04	3.12 min / 28.95 min = **0.11**

Notes: Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement.

The percentage of discussion time spent on responding to conceptual difficulties is shown in Table 4. Notably:

- The whole class discussions spent a *greater percentage* of time responding to student difficulties than the matched small group discussions in 3 out of 4 comparisons.
- The *total* amount of time spent on such discussion was longer in those whole class discussions also. In the medium level (HP) and lower level (CP) classes, the whole class discussions spent 4x and 6x as long, respectively, on addressing student conceptual difficulties as did the matched small groups.

Even for the most capable students observed, the AP classes, there did not appear to be any overall advantage for the small groups in having their conceptual difficulties addressed by discussion.

To What Extent Did Teachers and Students Support the Recognition, Use, and Interpretation of Key Visual Features of the Simulation?

Visual features identified in a pilot study as having played a key role were the movable Gravitational Potential Energy Reference Line and the Animated Energy Bar Graph (Figure 1). The two key concepts, the possibility that energy could take on negative values and the possibility that the total energy of a system could equal zero, could be explored by coordinated use of these two features. However, at times, we observed students experiencing difficulties with those two features—difficulties that appeared to be more perceptual than conceptual, where they misinterpreted the meaning of a feature or failed to find it at all. On the other hand, we observed small group students in some groups helping each other identify and use these interactive features.

Code: Student or teacher supports use and/or interpretation of a key visual feature or relationship in the simulation.* (*Here, by "interpretation of a feature," we mean the interpretation of its meaning, the development of some degree of understanding, as opposed to attaining rote knowledge or the ability to re-create a visual aspect through mimicry.)

This was coded when the student or teacher was observed engaged in one or more of the following to indicate or interpret a key visual feature or relationship: 1) **Selectively pointing out** some aspect of the visual feature or relationship; 2) **Giving a hint** to encourage use or interpretation of the feature; 3) **Gesturing in the air or over the display** to indicate this; 4) **Asking a question to prompt its use or interpretation;** 5) **Suggesting a manipulation of the simulation** to expose it; 6) **Pointing out a limitation** to interpreting its

meaning. Individual visual support 'moves' were identified and counted. The results below provide an estimate of what an individual student in the position of the camera could have been exposed to during each discussion.

Table 5: Support for key visual features (expressed as episodes / hour).

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	37 / 42.42 min = **52** per hour	4 / 23.90 min = 10 per hour
Yr 1 HP	Teacher A	26 / 62.03 min = **25** per hour	8 / 29.23 min = 16 per hour
Yr 1 AP	Teacher B	19 / 41.10 min = **28** per hour	10 / 32.32 min = 19 per hour
Yr 2 AP	Teacher B	19 / 41.71 min = **27** per hour	10 / 28.95 min = 21 per hour

Notes: Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement.

Frequencies of visual support episodes are given in Table 5.

- Rates of visual support episodes ranged from 10 to 21 per hour for the small group discussions and from 25 to 52 per hour for the whole class discussions.
- Total numbers of episodes ranged from 4 to 10 per small group discussion and from 19 to 37 episodes per whole class discussion.
- In no comparison did the small group discussion show an advantage.

Episodes of student-student support were included; it was not required that the person engaging in support be correct, only that the move appeared intended to help other students in addition to the supporter.

Activity Sheet Analysis

Did students recognize and use key visual features of the simulation?

We asked whether the activity sheets would give evidence for student recognition and use of key features. As students tried to describe their understanding of the concepts, they frequently mentioned the features or indicated them in drawings. The written and drawn answers to relevant open-ended questions were coded.

Code: Answer refers to GPE reference line in a way that implies that it is movable (as per a rubric).

Code: Answer contains evidence (as per a rubric) for use of at least one of 3 concepts supported by the Animated Energy Bar Graph having to do with changing energy and negative energy quantities.

Code: Answer contains evidence (as per a rubric) for use of a key relationship supported by coordinated use of the two key features (that TE and/or PE depend on position of the reference line).

Each student's answers were assigned a 1 or 0 for each code and an average was tabulated for the class section. Table 6 and Figure 2 show the percentage of students in each section whose work was assigned a 1 for a given code. In Table 6, whole class discussion data are listed above the matched small group data. Because the same data were scored along all three dimensions, the results are not added across dimensions.

Each group of 3 bars in Figure 2 represents a single class section analyzed along 3 binary dimensions listed in Table 6; each bar represents the percentage of students in that class exhibiting one of the three types of evidence. For instance, the first bar in each set represents the percentage of students who referred to the reference line in a way that implied it was movable. The light bars that would represent the CP and HP small group dimensions are not visible because they are zero for each of the three types of evidence. Notably:

- In each instance where the teacher facilitated whole class discussion about the relevant questions (did not inadvertently skip them; see Table 6), greater percentages of whole class students showed evidence in written and drawn work for having used the GPE reference line, for having used the Energy Bar graph, and for having used a key visual relationship between these two features.
- The only small group students who showed evidence on their activity sheets for having used the key features were Advanced Placement students. No student in the Honors Physics or College Preparatory small groups showed written or drawn evidence for having used either of the key features or the relationship between them.

These results, representing all students who answered the relevant questions on the activity sheets, suggest no advantage for the small group students over the whole class discussion students in making use of concepts supported by the key visual features, as evidenced in their written and drawn responses.

Table 6: Percentage of Students Showing Activity Sheet Evidence for Use of Key Visual Features

Class	Teacher	N	Lesson Format	Evidence for use of GPE ref line	Evidence for use of bar graph	Evidence for use of relationship
Yr 1 CP	Teacher B	11	WC	**0.36**	**0.27**	**0.18**
Yr 1 CP	Teacher B	13	SG	0.00	0.00	0.00
Yr 1 HP	Teacher A	20	WC	**0.10**	**0.05**	**0.05**
Yr 1 HP	Teacher A	18	SG	0.00	0.00	0.00
Yr 1 AP	Teacher B	13	WC	0.15*	0.23*	0.08*
Yr 1 AP	Teacher B	18	SG	**0.33**	**0.44**	**0.22**
Yr 2 AP	Teacher B	21	WC	**0.95**	**1.00**	**0.95**
Yr 2 AP	Teacher B	21	SG	0.81	0.95	0.48

Notes: Boldface indicates the larger mean scores within each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement. *Teacher inadvertently skipped the relevant questions on the activity sheet during whole class discussion.

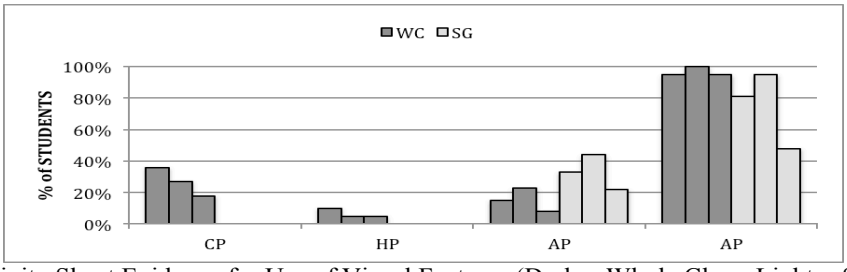


Figure 2. Activity Sheet Evidence for Use of Visual Features (Dark = Whole Class; Light = Small Group)

Discussion

Although almost all classes showed significant gains on the pre-post short answer questions, the teachers were surprised that there appeared to have been no pre-post advantage for students in the small group condition. This was true even though small group participants had had the advantage of hands-on experience with the simulations, opportunity for every student to raise questions with group-mates and with the teacher, opportunity for shyer students to speak up, and the engagement supported by small group work. Why did the small groups not do better than the whole class students? Our qualitative analyses suggest several plausible hypotheses.

First, even though the small groups had the advantage of hands-on experience with the simulations, our videotape analysis showed smaller percentage of *time spent on certain key concepts* than in the whole class discussions in 3 of 4 comparisons. Second, we identified a smaller percentage of *time spent on addressing student difficulties* in 3 of 4 comparisons. At times we observed student conceptual questions being ignored or only very briefly attended to in small groups, possibly due to a goal-oriented "complete the worksheet" mind-set (Conlin, et al., 2007). Third, 4 of 4 whole class discussions had more episodes (and greater frequency of episodes) where a teacher or student provided support for using the visual features of the simulations. As discussed in Stephens (2012), in small groups, even the episodes where students supported each other appeared to cluster around teacher visits to the group. Fourth, activity sheet analyses revealed that in 3 of 4 matched classes, students in the whole class discussions exhibited more evidence for actually having used the visual features in their own thinking. The only small group students who showed any evidence for this were in the highest-level classes in the study; none of the Honors or College prep small group students in these classes exhibited any written or drawn evidence for use of the visual features along any of the three dimensions examined. These results suggest the possibility, consistent with Wu and Huang (2007), that in certain situations there could be a disproportionate advantage for lower level students to participate in whole class discussion.

Conclusions and Implications

Small group students chose, on average, to spend less time on the activity sheets than was spent in whole class. We know this may have contributed to the preliminary quantitative pre post results, but we believe we have also identified other factors that can contribute. In this study, although each pair of classes was matched, the four pairs were not matched with each other in terms of student level and other factors. This means we are limited to within-pair comparisons as opposed to summing over the groups. Despite these small samples, we believe that the four pairwise comparison studies are sufficient to raise questions about the common assumption that small groups are always a better configuration, and suggest hypotheses as to why they may not be.

Our classroom observations suggest that teachers may need more guidance provided along with simulations to help them identify what features and relationships are likely to be overlooked by students;

teachers may also need suggestions for making these features explicit. These results appear to offer encouragement to teachers who do not have the resources to allow their classes to engage regularly in small group work at the computer. The argument here is not that the small group work did not have benefits—it clearly did; small group students had pre-post gains almost as, if not as, large as the whole class students. Rather, we argue that the whole class and small group formats could have had compensating strengths and weaknesses when it came to learning from the sophisticated physics simulation used here. The fact that the students in these whole class discussions matched or exceeded the performance of their small group peers implies that whole class strategies evidently exist that can promote at least some of the active thinking and exploration that has been considered to be the strength of hands-on small group work. We suggest that a mixture of the two formats might be optimal; further investigation is warranted to see which might be best used when.

Most of the teachers in our studies believed that simulations are much more effective when used by small groups; however, preliminary pre-post analysis did not appear to bear this out. In order to investigate what was happening in class discussions in the two formats, we conducted qualitative analyses of matched whole class and small group discussions that accompanied use of an interactive physics simulation. These revealed that in the four whole class discussions, there was generally 1) more time spent on key concepts; 2) more time addressing student conceptual difficulties; 3) more episodes providing support for using key visual features of the simulations; 4) more evidence for student use of the visual features in their writing and drawing. These results are consistent with similar analysis of a different lesson sequence (Stephens & Clement, in preparation). Our results suggest the possibility that there may be certain instructional situations where there is an advantage to spending at least part of the time with an interactive simulation in a whole class discussion mode.

References

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman, C. E. (2008). A study of educational simulations Part I - Engagement and learning. *Journal of Interactive Learning Research, 19*(3), 397-419.

Bennett, J., Hogarth, S., Lubben, F., Campbell, B., & Robinson, A. (2010). Talking science: The research evidence on the use of small group discussions in science teaching. *International Journal of Science Education, 32* (1), 69-95.

Conlin, L., Gupta, A., Scherr, R., & Hammer, D. (2007). The dynamics of students' behaviors and reasoning during collaborative physics tutorial sessions. *AIP Conference Proceedings*, *951*(1), 69-72.

Cook, M. (2006). Visual Representations in Science Education: The Influence of Prior Knowledge and Cognitive Load Theory on Instructional Design Principles. *Science Education 90,* 1073-1091.

Linn, M. (2003). Technology and science education: Starting points, research programs, and trends. *International Journal of Science Education, 25*(6), 727–758.

Mayer, R. E., and Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390-397.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher, 44*, 18-23.

Raghavan, K., Sartoris, M., & Glaser, R. (1998). Why does it go up? The impact of the MARS curriculum as revealed through changes in student explanations of a helium balloon. *Journal of Research in Science Teaching, 35*(5), 547-567.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning, 19*, 9– 20.

Russell, J., & Kozma, R. (2005). Assessing learning from the use of multimedia chemical visualization software. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 229-332). Dordrecht, The Netherlands: Springer.

Stephens, L. (2012). *Student recognition of visual affordances: Supporting use of physics simulations in whole class and small group settings*. (Doctoral Dissertation). University of Massachusetts, Amherst. *Open Access Dissertations*. http://scholarworks.umass.edu/open_access_dissertations/668.

Stephens, L. & Clement, J. (in preparation).

Williams, M., Linn, C., Ammon, P., & Gearhart, M. (2004). Learning to teach inquiry science in a technology-based environment: A case study. *Journal of Science Education and Technology, 13*(2), 189-206.

Wu, H.-K., & Huang, Y.-L. (2007). Ninth grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. Science Education, 91, 727-749.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants DRL-1222709 and DRL-0723709 awarded to John J. Clement. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The Contours and Possibilities of Desire in Sociocultural Research on Learning and Becoming

Ian Parker Renga, University of Colorado at Boulder, 249 UCB, Boulder, CO, 80309, ian.renga@colorado.edu

Abstract: Though it has enhanced the learning sciences in numerous ways, sociocultural theory has only somewhat dealt with desire. There is some indication that desire is learned in social activity and informs identity development, which whets the appetite for conceptual elaboration. In this paper I draw from scholarship in philosophy, religious studies, and education to define and expand the concept. This work reveals how desire is directed, disciplined, and has consequences of critical concern. Conceptually, desire may also be preferable to motivation for researching and appreciating what Dorothy Holland (1992) labeled the *directive force* evident in individuals' participation within social activity systems. I refer to my own work in teacher education to illustrate how the concept frames important questions and lines of inquiry for further research.

When investigating learning within social activity systems, what are we to make of participants' expressed yearnings, motivations, aspirations, pleasures, or even dreams? Indeed, I find that an intriguing aspect of learning—*desire*—still remains underdeveloped from a sociocultural perspective. Compelling cases have been made that desire is both socially constructed (Holland, 1992) and part of identity formation (Packer & Goicoechea, 2000), suggesting that what we come to yearn for may profoundly influence the kind of people we are and are becoming, and what we do and how we see ourselves may shape the objects and intensity of our yearning. As I propose in this paper, desire is a concept that can enrich our understanding of the relationship between practice, identity, and *telos*. But just as much of our learning is ubiquitous and tends to go unnoticed (Lave, 2000), desire is likely present but unappreciated in daily activity. As I will reveal, desire is an evocative concept that can open up emotional terrain that tends to elude characterization. This may be its conceptual virtue in an era driven by post-Enlightenment rationality where feelings of desire are often treated as threats to order and stability—irrational urges to be subordinated to reason and either mastered or driven into hiding, accessible only through religion or the arts.

In this paper I argue that sociocultural theory has much to offer our understanding of desire, and the concept may prove useful to research in the learning sciences. I begin by drawing upon scholarship in philosophy, religious studies, education, and critical theory to illuminate some of desire's conceptual contours and dilemmas. From there I locate it within sociocultural theory and research. I conclude by briefly showing how desire is informing the direction of my study of teacher learning within teacher preparation programs.

Conceptual Roots and Theories

The term 'desire' comes from the Latin *desiderare* (to long for), which in turn comes from *de sidere* (of the stars) (Hollis, 2010). It would appear that our language of yearning has roots in early travelers' use of the stars to mark distance and direction. Long before Facebook or Google Maps, a glance upward to the cosmos could stir wanderlust or remind weary travelers of the comforts of far off homelands. The ancients also ascribed divine meaning to the stars, projecting onto them mythic ideals and purposes for human activity. While our mythic imagination may have waned in modern times (Armstrong, 2009), we persist in conjuring up and fervently chasing our dreams and ideals. These dreams serve an important function as a means for playing out potential futures that have yet to materialize (Simon, 1992). They situate the motion of life between points past and future, actual and possible, thereby establishing various trajectories of being (Cole, 1998; Polman & Miller, 2010). They also draw attention to our relative velocity and what we require to keep moving or keep up with others. Indeed, talking with our fellow travelers, we might wonder how our desires compare and why we find certain futures more compelling than others.

Throughout history philosophers have sought to understand desire and its function in the activity of daily life. Characterizing this large body of work, Timothy Schroeder (2009) suggests that desire is fundamentally a state of mind concerned with fulfilling a want or perceived need. He identifies several major theoretical families that have formed around various interpretations of desire's presence and purpose in our lives. *Action-based* theories define desire as dispositional, or the strength of one's inclination to take whatever action is believed will fulfill one's yearning. This tends to shortchange emotional features of having and acting on desire, which has given rise to *pleasure-based* theories focusing on the compulsion to seek fulfillment and satisfaction. Other philosophers have argued that this elevates hedonistic impulses and is dismissive of the human propensity for moral evaluation. They maintain that desire is *good-based*, which follows from the Socratic notion that people desire that which they think is good. Still others prefer *attention-based* theories of desire based on the tendency for individuals to fixate on some inclination around which they construct reasons

to satiate the fixation. Finally, *holistic theories* of desire incorporate some admixture of pleasure, morality, and attention, which are seen as overlapping and often inseparable.

Schroeder also highlights additional dilemmas that appear in theoretical discussions of desire. One is making sense of multiple desires. Indeed, he notes that humans have many different desires and expressions of desire that feed into one another. Imagine a person standing on the shore of a cool lake on a hot day: that person can feel the urge to jump into the lake (a desire for pleasure) and also want to be the kind of person who acts on such impulses (a desire for recognition). Another challenge is what to make of relative intensity, or situations where a desire is erratically enacted despite its consistent and unwavering presences in a person's life. For example, teachers often have bad teaching days, but that does not necessarily signal a faltering desire to ensure their students' welfare. Also, certain desires may increase or decrease in intensity with respect to circumstance and social conditions.

Schroeder (2004) helpfully distills these theories into three major components that he argues should be examined in any exploration of desire: pleasure (and displeasure), motivation (and inhibition), and reward (and punishment). *Pleasure* is evident in a sense of satisfaction, specifically the experience of having our expectations at least met, and preferably exceeded, within a given situation. *Motivation* is the conscious focus on goal attainment—the fixation on a particular outcome and how persistent a person is to attain it. And *reward* is the feedback or stimulus received in response to activity. As I interpret them, each component evokes a suite of reflective questions:

- *Pleasure*: What do I require to be satisfied? What do I enjoy about an activity I am engaged in? How does it compare to my expectations?
- *Motivation*: What preoccupies my mind? What am I after? Why do I choose to take certain actions? What do I hope to accomplish through those actions?
- *Reward*: How did it feel to take certain actions? What feedback keeps me pursuing my goals? Why do I avoid some situations but seek out others?

How we answer these questions may reveal how our desires permeate, shape, and direct who we are and what we do. Schroeder observes that the components overlap and inform one another. In fact, he argues that reward is underappreciated in contemporary discussions of desire because it is seen as invoking the ghost of behaviorism. And yet, he contends that acknowledging rewards and feedback is crucial for understanding how desire is learned, and he points to recent neurobiology research on the brain's reward centers to support his claim. Whether or not one buys the biological argument (for a dissenting view, see Latham, 2006), Schroeder's suggestions that desire is both experienced bodily and is learned have important implications for how the concept is framed and studied, especially when examined through critical and sociocultural lenses. Before taking up this examination, I want to offer additional conceptual insights from the humanities.

Desire as Directed and Disciplined

Discussions of desire in the humanities tend to highlight three key conceptual features: 1) the object toward which desire is directed, 2) the ways in which desire is focused and disciplined, and 3) the intended and unintended consequences of enacting desire. For example, in his assessment of whether or not sports develop character, Mark Edmundson (2013) refers to Plato's notion that human's possess an inherent desire for glory that necessitates reason to control it. Edmundson concedes that sports can offer an outlet for this desire by permitting a disciplined pursuit of the object—glory—without violating the rule of law and moral sensibilities. But he argues that the positive consequences of playing sports are debatable, as deeper and more insidious desires for power and individual recognition often lie hidden beneath the cloak of concern for character development.

Religious scholars and writers inquiring into the human condition have much to say on the nature of desire. In his fourth century text, *Confessions,* St. Augustine chronicles his efforts to discipline and direct his desire for pleasure. He recalls the sexual exploits of his youth and how he eventually tired of his insatiable appetites and turned to God for comfort and direction. Exploring this text, Margaret Miles (1992) observes how Augustine presents a contradictory framing of pleasure. On the one hand he sees pleasure as something worth maximizing through sustained intensity. But on the other he admits that perpetual bliss is impossible, which his own story makes clear. Augustine resolves the contradiction by vowing to only desire God, in whom he believes pleasure is readily and endlessly available without the consequences of guilt or hangovers. He thereby finds an object for his desire that is suitably beyond his comprehension yet compelling enough to warrant his devotion. Thoreau (1854/2001) reaches a similar conclusion in *Walden* while pondering his experiences in nature. He expresses a firm Puritanical belief that hard work develops divine character, suggesting that such character is more easily maintained when the body is too exhausted to be lustful. Like Augustine, he argues that desire must be controlled and directed toward a higher ideal, a purity of being focused on an incomprehensible natural beauty. He worries that his more earthly desires will spoil his devotion to nature and its Creator. For Thoreau the solution is clear: turn one's attention to God so "the spirit can for the time control every member and function of the body, and transmute what in form is the grossest sensuality into purity and devotion" (p. 514).

The challenge for contemporary devotees is disciplining one's desire within a consumer-driven culture. For Vincent Miller (2004) "[i]t is clear that our desires are shaped, encouraged, and manipulated" by our daily participation in a society fixated on consuming (p. 109). He sees the formation of this desire as happening through two devices: misdirection and seduction. *Misdirection* is a tactic whereby individuals become convinced that they can fulfill their innermost desires—those reflecting deeply held values and beliefs—through their consumptive activity. Savvy merchants cleverly associate the purchase of their products with popular causes, such as cancer research and environmentalism. *Seduction* keeps individuals in a state of perpetual hunger and convinces them that the ideal state of being is one of constant dissatisfaction. Elaborating this point, William Cavanaugh (2008) observes that consumer culture focuses attention on choice and cultivates a sense of pleasure in the yearning for things. Happiness is having options, leafing through catalogues and combing through websites. The moment of decision—of actual consumption—feels oddly unsatisfying, thereby initiating the drive to pursue other desires and choices.

Miller contends that seduction and misdirection generate a stunted and dispersed form of desire that promises much but offers little. He argues that from childhood onward we learn to repress our desires and engage in consumptive activity to relieve the "ocean of desire seething beneath what can be directly said or demanded" (p. 125). But consuming fails to address the depths of that ocean. It merely skims the surface, converting the rich cultural textures of human activity and meaning into comparable units—commodities—that diminish their possible effects on our lives. Such cultural commodification leads to a shallow appropriation of cultural and spiritual symbols, languages, and identities. More significantly, it deludes individuals into believing that symbolic possession can bring about cultural and spiritual ideals that require thoughtful and sustained work. When symbolic consumption inevitably fails to prove transformative, people come to see such ideals as illusions, and visions of happiness, peace, justice, and the fullness of God are written off as utopian fantasies. The yearning for such ideals thus wanes, and the resulting vacuum is filled with a "closed, disenchanted universe… of technological planning and pragmatic rationality in which expectation plays no role" (p. 131). Living is thereby reduced to seemingly endless oscillations between cool, levelheaded practicality and hot, impulsive consumer indulgence. Miller contends that breaking this pattern and realizing broader ideals requires vocational practice within communities that nurture a "commitment to a particular form of life and to the transformation of the self in order to sustain that commitment" (p. 137).

James Smith (2009) suggests that Christian educators interested in cultivating vocation should focus less on transmitting a worldview and more on directing desire. To this end he poses the following questions:

> What if education … is not primarily about the absorption of ideas and information, but about the formation of the hearts and desires? … What if education was primarily concerned with shaping our hopes and passions—our visions of "the good life"—and not merely about the dissemination of data and information as inputs to our thinking? What if the primary work of education was the transforming of our imagination rather than the saturation of our intellect? And what if this had as much to do with our bodies as with our minds? (p. 18)

As with sociocultural learning theorists, Smith challenges the view that an education is primarily a cerebral affair—a matter of belief, view, or knowledge. He also places a premium on learning environments and sees a strong connection between teaching and ontology, noting how "behind every constellation of educational practices is a set of assumptions about the nature of human persons—about the kinds of creatures we are" (p. 28).

Secular scholars of education have similarly found purchase in linking teaching to desired futures and ways of being. Jacqueline Cossentino (2005, 2006) has documented how the Montessori teaching tradition has deep roots in a specific cosmic vision of the good life that gives meaning and purpose to its practices and practitioners. And Jim Garrison (1997) notes how Dewey framed teaching as fundamentally driven by educational commitments to particular values and visions of society. Garrison contends that great teaching requires educators to directly engage these visions and the often intensely felt passions they foment. This is why, he observes, the ancient Greeks saw romantic desire—*eros*—as vitally important to education. What stars to follow and why were considered significant matters because "[w]e become what we love. Our destiny is in our desires, yet what we seek to possess soon comes to possess us in thought, feeling, and action" (p. xiii).

Dan Liston (2004) sees such passion as the lure of learning—*eros* manifested in a desire to experience the awe and wonder of the 'grace of great things'. For teachers this desire can find expression in efforts to connect children to the world and the various ways it can be understood and experienced. Christine Downing (2009) gives voice to this sentiment when she says of her own teaching,

> …I want to communicate my love—not exactly of my students, though not exactly *not* of my students—but more explicitly my love of the books, the authors, the ways of looking at the world that have moved and inspired me, and my love of the process of inquiry that brought me

> to those books, those writers; I want to encourage my students to find what might move them in the same way—and come to love looking for it. (italics in original, p. 153)

Mark Edmundson (2005) expresses a similar fondness for books and their usefulness for deciding how and why to live one's life in particular ways. By situating the activity of reading within a dynamic exchange between love and identity, both Downing and Edmundson acknowledge the power of narrative to inform passion and purpose. Liston (2004) suggests that without desire teaching and learning are lifeless endeavors, and students are unlikely to inherit a love of learning. He observes that teachers who build their practice upon a love of learning maintain a heartfelt connection to their work. They come to appreciate how teaching can be an artful instruction not only in ways to acquire knowledge but also in how to love and cherish that knowledge in particular ways.

Despite its promise, an engagement with passionate desire is arguably rare in contemporary U.S. schools. Jennifer Logue (2012) suggests that this is because rosy portrayals of educational *eros* tend to overlook the darker side of desire. Less discussed is how desire can drive people to defensiveness and obsessive, hysterical behavior. Indeed, in teaching, unfulfilled desire can be deeply disheartening (Liston, 2000). Logue suggests that educational institutions are designed to prevent desire's ill effects. They require order and predictability to function smoothly, which breeds a preference for more sanitized and measurable forms of learning. Erotic desire threatens this order as a force that upends certainties while feeding radical possibilities. According to Logue, "The unruly, unpredictable, and unconscious aspects of eros render the school or university's attempt to mold and monitor a knowable, disembodied self, a self measured by predetermined skills and predictable outcomes, impossible" (p. 73). As such, *eros* is often confined or scrubbed from formal educational settings, though not without costs.

Desire through a Critical Lens

Within the activities of directing, disciplining, and instructing desire are issues of privilege and power that warrant critical interrogation. I will briefly offer a feminist examination of desire, though critical race and class analyses would be illuminating and are certainly necessary to appreciate the lived experiences of the concept.

Judith Butler (2004) maintains that our desires are interwoven with the norms of enduring gender narratives. Because these narratives originate externally to us, we can assume that the desires coded within them also originate externally. Miles (1992) similarly observes how "desire is always marked by the particularities of individual lives, by socially constructed gender assumptions, expectations, and roles, by social location, institutional affiliation, class, and race" (p. 135). She points out that Augustine's unchecked privilege as a heterosexual male leads him to define desire as something both egocentric and ravenous, which in turn leads him to prescribe the authoritative disciplining of religious practice. Miles observes how, in his writing, Augustine builds excitement with tantalizing and titillating details only to switch suddenly into chastisement and the necessity of finding fulfillment in God. He thus frames a choice between desiring earthly objects and desiring God—a choice that was and arguably still is a privilege of men, whose desires are publically sanctioned. This indulgence/repentance tension is problematic from a critical perspective in that it permits men, and those in privileged positions of power more generally, to excuse their excesses and establish their pleasure as something only God has the authority judge.

Despite these critiques, Miles finds it unfair to hold Augustine accountable for assumptions and transgressions exceeding the social consciousness of his time, especially if it leads to the outright rejection of his work. For Miles, *Confessions* still offers keen and arguably timeless insights about the human condition. Even so, she advises greater caution when reading it than has been typical of religious scholarship over the centuries. This body of work has helped establish the dominant norms and conditions for gendered identity, with men as the desiring beings and women the objects of their desire. Miles notes the conspicuous absence of the voices of women in Augustine's text; without these perspectives we cannot fully appreciate the constructive moments between friends and lovers where desire is formed. Nor can we understand what the women made of the situation. How did they participate in a masculine narrative where women are considered pure, their earthly desires severely limited and controlled? We can reasonably assume that it was difficult, and probably joyless. Even today young women are denied a publically acceptable means for exploring embodied desire and crafting certain identities and possible futures (Fine & McClelland, 2006; Tolman, 2012). According to Butler (2004), any woman or person identifying as gay, lesbian, or transgender who challenges these arrangements is forced to ask

> If I desire in certain ways, will I be able to live? Will there be a place in my life, and will it be recognizable to the others upon whom I depend for social existence? (p. 3)

Butler maintains that both desire and social norms are important features of human thriving and civility. As such, the conscientious resistor should articulate a new vision of being and desire and the necessary conditions for its enactment. According to Michelle Fine and Sara McClelland (2006), this requires a more

holistic view of desire—a "thick desire"—that "situates sexual well being within structural contexts that enable economic, educational, social, and psychological health" (p. 301). The liberalization of the modern economy, with its promise of personal freedom and self-formation, would seem more conducive to a thicker form of critical, embodied desire. While arguably an improvement, Miller (2004), evoking Foucault, contends that the liberal shift is somewhat illusory, as the injustice of confessional monitoring has given way to the injustice of self-monitoring. In a liberal society, individuals are expected to proclaim their desires and accept full responsibility for their misdirected desires. Significantly, this leaves externalities and social conditions out of the conversation. But such conditions are indeed present and serve as crucial resources for constructing desire, which a sociocultural lens helpfully spotlights.

Desire from a Sociocultural Perspective

Sociocultural theory dissolves firm distinctions between individuals and social conditions. Through its development, the theory has challenged the prevailing model of learning as acquisition and transfer (Sfard, 1998; Rogoff, 1997; Packer, 2001) giving way to conceptions of learning as social activity situated within communities (Lave & Wenger, 1991; Wenger, 1998), undertaken in complex figured worlds (Holland, Lachiotte Jr., Skinner, & Cain, 1998), profoundly influenced by history and culture (Gutiérrez & Rogoff, 2003), and ultimately organized toward possible futures and identities (Cole, 1998; O'Connor & Allen, 2010; Polman & Miller, 2010). In formal educational settings, researchers have applied the theory to upend problematic views of students, teaching, and learning. Hand-in-hand with a postmodern, critical perspective, sociocultural theory has developed with a desired outcome (*telos*) in mind of a just and equitable society (O'Connor & Penuel, 2010).

By treating individuals as situated within and constitutive of social conditions, the theory offers ample conceptual groundwork for observing and interpreting how desire takes shape and is learned. It provides helpful tools for empirically analyzing questions such as *How do our desires form?*; *How do we come to prioritize some 'objects' of desire, including some imagined futures, over others?*; and *How do we come to understand and pursue our desires in certain ways?* I would point out that these questions and their underlying concerns have been raised in the sociocultural literature by the mention of a drive to take action or participate in communities of practice. In the first paragraph of their seminal text, Lave & Wenger (1991) refer to the novice's "*intentions* to learn" as initiating their engagement with a particular community of practice (italics mine; p. 29). And Barbara Rogoff (1997) suggests that evidence of learning can include a person's "changing purposes for being involved, commitment to the endeavor, and trust of unknown aspects of it (including its future)" (p. 280). We might therefore ask how and in what ways these intentions and purposes form and influence participation within social activity systems.

Doing so from a socioculturally informed perspective provides conceptual inroads into aspects of learning that have been predominantly defined by cognitive psychology in terms of motivation, which is treated as something that resides within individuals (cf., Stipek, 1993) to be altered through intrinsic or extrinsic means (cf., Ginsberg, 2005; Middleton, 1995). In their sociocultural take on student motivation, Robert Rueda and Luis Moll (1994) challenge this view and define motivation as situated in social interactions within specific cultural contexts. They note how this reframing permits a more dynamic understanding of individual student participation and its relationship to the (de)motivating conditions of classrooms. They argue that such a view can aide teachers in personalizing the classroom learning experience such that it cultivates students' motivations.

Instead of an exclusive focus on motivation, Dorothy Holland (1992) opts to study desire in her investigation of college women's pursuit of romance. She refers to desire as a *directive force* that is not natural but cultural, something learned through social participation. Holland maintains that "thoughts and feeling, will and motivation, are *formed* as the individual develops" (italics in original, p. 63). This formation is accomplished through social activity and engagement with cultural resources, and it serves to direct and discipline one's thoughts and feelings in particular ways for particular purposes. For the women she studied this involved learning various discursive moves from peers to deflect unwanted attentions from some suitors while eliciting attentions from others. Holland observed how the resulting world of romance was formed and reinforced by the women's behavior, with rules for participation and penalties for violations. Living within this world directed the women toward particular outcomes and a particular vision of the good life. Through their participation in this world, the women came to forge an identity rooted in some desired futures but not others, with potentially harmful implications for their academic participation and professional opportunities in college and beyond (Holland & Eisenhart, 1990).

Holland's decision to focus on desire instead of motivation is noteworthy. Of course, evoking romantic love—*eros*—invites entry into a family of concepts that includes passion and desire. Significantly, though, Holland did not treat romance as a factor affecting some inherent conception of motivation, but rather positioned romance as an organizing object giving shape and purpose to the women's social activity and identities. Studies of motivation, even when socioculturally framed, tend to take motivation rather than its various targets as the primary focus of inquiry. Holland organizes her study instead around understanding the object of the women's

interest—romance—and how it takes hold within a social activity system. Thus, so-called individual motivation is not only situated socially, it is situated within the future-oriented, future-creating character of participation within communities. As I see it, acknowledging this shifts the focus away from assessing individuals' relative motivational strength and draws attention to conflicts in desired future outcomes, such as achieving romantic fulfillment versus a high-status professional identity. Those futures, while crafted through and from the experiences of sociocultural, embodied activity, exist in the imagination. And, as Schroeder (2004) points out, it is possible for people to do nothing to realize imagined futures—to show no clear signs of being motivated—yet still desire them.

Holland's sociocultural treatment of desire arguably passes on the modernist tendency to presume foundational origins of human activity in favor of a postmodern understanding of activity and its drivers as socially constructed with respect to imagined futures (Packer, 2000). Doing so offers insights into objects of sociopolitical importance formed within and across communities, like identity (Urrieta Jr., 2007). Martin Packer and Jessie Goicoechea (2000) assert that identity is a construct formed through social participation "in relationships of recognition and desire" (p. 228). They note how participation is complicated by power, politics, and the fact that people tend to move between communities. They maintain that, as individuals try to make sense of this movement and their participation, a desire for recognition emerges and drives the quest for a meaningful identity. As they explain, this "desire directed toward another person … seeks recognition that gives not just consciousness of self but self-consciousness" (p. 233). Thus, we do more than form an identity in relationship to others; we form the desire for an identity and the desire that others acknowledge it.

Framing a Research Agenda

Through a sociocultural framework, desire can invigorate familiar conceptual terrain by highlighting the direction of participation, practice, and identity formation. It does so by invoking the *why* questions encoded within social activity systems: Why is action taken? Why are some goals elevated over others? Why do individuals seek to become certain kinds of people? It encourages inquiries into the imagined futures and directions, whether intended or not, of social activity and how individuals are helping to establish those futures and directions and ensure that progress is made. It also permits us to take seriously not just emotions, but experiences of pleasure.

In my research on teacher education I am using the conceptual language of desire to understand teacher learning as framed by reformers arguments for *what* works in schools and *who* should be tapped to do the work. In the U.S. there is a growing call for a practice-based approach to teacher training and assessment that reformers believe will bolster the teaching profession by grounding it in research-supported best practices (Zeichner, 2012). There has been a concurrent push to identify inherent traits and desires for the work of teaching to guide teacher selection and training (Richardson & Watt, 2006; Watt & Richardson, 2008). To this end, personality instruments like the TeacherFit Inventory are being used to identify teaching applicants by how well their characteristics match those considered necessary for the work. Some practice-based reformers contend that such efforts are misguided and distract from the work of identifying effective teaching practices (Hiebert & Morris, 2012). They maintain that nobody is born a teacher and the practices of effective teaching must be learned (Ball & Forzani, 2009). Of course, the prospective teacher must be somebody who is willing to learn those practices and accept certain definitions of effective. Stated differently, they need to be comfortable with developing particular desires for the work of teaching.

Indeed, seen through the lens of desire, any educational reform proposal can be interpreted as an effort to set the conditions for job satisfaction, directing teachers' attention to particular goals for their work and establishing what counts as worthwhile feedback. That some teachers may be turned off by these conditions while others find them invigorating could be interpreted as a matter of fit between individuals and the work of teaching. But seeing such reactions as a matter of differing desires for teaching—for what it is, does, and should accomplish—encourages a more nuanced inquiry into those conditions and what drives and sustains them. It begs questions such as: Why are certain experiences of job satisfaction considered more acceptable than others? Why are some goals deemed more worthy of pursuit than others? Why are some kinds of feedback offered over others? Pointing out purported school realities to answer such questions may be inadequate. Cleo Cherryholmes (1992), referring to Dewey's pragmatic critique of a fixed social reality, notes, "Not everything that works is desirable, not every belief that is 'true' is to be acted upon" (p. 14). The error may be thinking that "true beliefs" about teaching only reflect school realities rather than construct them (Berger & Luckmann, 1966/1991). As a consequence, what is desired of teaching may become limited by self-fulfilling ideas about what it can accomplish, while dreams of what it could (or should) accomplish are cast aside and left unexplored.

In an effort to understand desire in teacher learning, I am currently employing grounded theory methodology (Charmaz, 1995; Charmaz & Belgrave, 2003) to investigate beginning teachers' participation within their teacher education programs. Similar to Holland (1992), I want to understand how certain objects of desire, including imagined futures, emerge from and give shape to a social activity system and individuals' participation within it. I am curious to know what the beginning teachers yearn for—what they come to desire

and how they pursue their desires—as they interact with one another, teach children, talk with mentors, receive feedback, and engage programmatic practices, norms, and rituals.

Conclusion

In this paper I have merely scratched the surface of the many ways desire is discussed and can illuminate the transformative texture of learning. To explore its conceptual features I have turned to humanities scholarship concerned with desire's role in crafting the kinds of people we are and are becoming. I have also brought in a critical view to interrogate a particular narrative of desire that limits its expression and enactment. I have suggested that desire is often overlooked in daily activity, its visibility likely requiring engagement with the imagination and the wishes, dreams, and futures developed therein. Just as sociocultural theorists have constructed definitions of learning for specific purposes, I want to establish desire as a conceptual construct based on sociocultural assumptions about the nature of human activity such that various expressions and experiences of yearning are analytically accessible in research on teacher learning. Humanities scholarship and studies such as Holland's have laid the groundwork for investigating desire's presence and consequences within specific contexts of social activity, like teaching and teacher education. This paper represents a next step in conceptualizing desire, and empirical research is needed to further establish its contours and analytic usefulness within the learning sciences.

References

Armstrong, K. (2010). *The case for God*. New York: Alfred A. Knopf.

Ball, D. & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497-511.

Berger, P.L., & Luckmann, T. (1966/1991). *The social construction of reality: A treatise in the sociology of knowledge* (No. 10). London, UK: Penguin.

Butler, J. (2004). *Undoing gender*. New York: Routledge.

Cavanaugh, W.T. (2008). *Being consumed: Economics and Christian desire*. Grand Rapids, MI: Eerdmans Publishing.

Charmaz, K. (1995). Grounded theory: Rethinking methods in psychology. In J.A. Smith, R. Harré, and L. vanLangenhove (Eds.) *Rethinking methods in psychology*. Thousand Oaks, CA: Sage Publications.

Charmaz, K., & Belgrave, L. (2003). Qualitative interviewing and grounded theory analysis. In J.F. Gubrium, J.A. Holstein, A.B. Marvasti, and K.D. McKinney (Eds.) *The SAGE handbook of interview research: The complexity of the craft*. Thousand Oaks, CA: Sage Publications.

Cherryholmes, C.H. (1992). Notes on pragmatism and scientific realism. *Educational researcher*, *21*(6), 13-17.

Cole, M. (1998). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.

Cossentino, J.M. (2005). Ritualizing expertise: A non-Montessorian view of the Montessori method. *American Journal of Education*, *111*(2), 211-244.

Cossentino, J.M. (2006). Big work: goodness, vocation, and engagement in the Montessori method. *Curriculum Inquiry*, *36*(1), 63-92.

Downing, C. (2009). The eros of teaching. In Slattery, D.P., and Selig, J.L. (Eds.) *Reimagining education: Essays on reviving the soul of learning*. New Orleans, LA: Spring Journal.

Edmundson, M. (2005). *Why read?* New York: Bloomsbury.

Edmundson, M. (2013). Do sports build character? In *Why teach?: In defense of a real education*. New York: Bloomsbury.

Fine, M., & McClelland, S.I. (2006). Sexuality education and desire: Still missing after all these years. *Harvard Educational Review*, *76*(3), 297-338.

Garrison, J. (1997). *Dewey and eros: Wisdom and desire in the art of teaching*. New York: Teachers College Press.

Ginsberg, M.B. (2005). Cultural diversity, motivation, and differentiation. *Theory into practice*, *44*(3), 218-225.

Gutiérrez, K.D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational researcher*, *32*(5), 19-25.

Hiebert, J., and Morris, A.K. (2012). Teaching, rather than teachers, as a pathway to improving classroom instruction. *Journal of Teacher Education 63(92),* 92-102.

Holland, D. (1992). How cultural systems become desire: A case study of American romance. In R. D'Andrade & C. Strauss (Eds.), *Human motives and cultural models* (pp. 61-89).

Holland, D.C., & Eisenhart, M.A. (1990). *Educated in romance: Women, achievement, and college culture*. Chicago, IL: University of Chicago Press.

Holland, D., Lachicotte Jr.,W., Skinner, D., & Cain, C. (1998). *Identity and agency in cultural worlds.* Cambridge, MA: Harvard University Press.

Hollis, J. (2010). Permutations of desire. *Parabola 35 (3),* p. 6-11.

Latham, N. (2006). Three compatible theories of desire. *Dialogue* 45, 131–138.
Lave, J. (2000). The practice of learning. In Knud Illeris (Ed.) *Contemporary theories of learning: Learning theorists… in their own words*. New York: Routledge.
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
Liston, D.P. (2000). Love and despair in teaching. *Educational Theory*, *50*(1), 81-102.
Liston, D.P. (2004). The lure of learning in teaching. *The Teachers College Record*, *106*(3), 459-486.
Logue, J. (2012). Erotic study and the difficulties of desire in education. *Philosophy of Education Archive*, 72-75.
Middleton, J.A. (1995). A study of intrinsic motivation in the mathematics classroom: A personal constructs approach. *Journal for Research in Mathematics Education*, 254-279.
Miles, M.R. (1992). *Desire and delight: A new reading of Augustine's* Confessions (p. 23). New York: Crossroad Publishing.
Miller, V.J. (2004). *Consuming religion: Christian faith and practice in a consumer culture.* New York: Continuum.
O'Connor, K., & Allen, A. (2010). Learning as the organizing of social futures. *NSSE Yearbook*, *109*(1).
O'Connor, K., & Penuel, W.R. (2010). Introduction: Principles of a human sciences approach to research on learning. *NSSE Yearbook, 109(1)*.
Packer, M.J. (2000). Escaping modernity. *Human Development*, *43*(3), 170-177.
Packer, M.J. (2001). The problem of transfer, and the sociocultural critique of schooling. *Journal of the Learning Sciences*, 10(4), 493-514.
Packer, M.J., & Goicoechea, J. (2000). Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educational Psychologist, 35(4),* 227-241.
Polman, J. L., & Miller, D. (2010). Changing stories trajectories of identification among African American youth in a science outreach apprenticeship. *American Educational Research Journal*, *47*(4), 879-918.
Richardson, P.W., & Watt, H.M. (2006). Who chooses teaching and why? Profiling characteristics and motivations across three Australian universities. *Asia-Pacific Journal of Teacher Education, 34(1)*, 27-56.
Rogoff, B. (1997). Evaluating development in the process of participation: Theory, methods, and practice building on each other. *Change and development: Issues of theory, method, and application*, 265-285.
Rueda, R., & Moll, L.C. (1994). A sociocultural perspective on motivation. *Motivation: Theory and research*, 117-137.
Schroeder, T. (2004). *Three faces of desire*. New York: Oxford University Press.
Schroeder, T. (Winter 2009). "Desire". *The Stanford Encyclopedia of Philosophy.* Edward N. Zalta (Ed.). Retrieved at http://plato.stanford.edu/archives/win2009/entries/desire/
Sfard, A. (1998). On two metaphors for learning and the danger of choosing just one. *Educational Researcher 27(2),* 4-13.
Simon, R. (1992). *Teaching against the grain: Texts for a pedagogy of possibility*. New York: Bergin & Garvey.
Smith, J.K.A. (2009). *Desiring the kingdom: Worship, worldview, and cultural formation*. Grand Rapids, MI: Baker Academic.
Stipek, D.J. (1993). *Motivation to learn: From theory to practice*. Needham Heights, MA: Allyn and Bacon.
Thoreau, H.D. (1854/2001). *Walden*. In *The works of Henry David Thoreau*. Ann Arbor: State Street Press.
Tolman, D.L. (2012). Female adolescents, sexual empowerment and desire: A missing discourse of gender inequity. *Sex Roles*, *66*(11-12), 746-757.
Urrieta Jr., L. (2007). Figured worlds and education: An introduction to the special issue. *The Urban Review*, *39*(2), 107-116.
Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker*, *9*(5), 2-3.
Watt, H.M., & Richardson, P.W. (2008). Motivations, perceptions, and aspirations concerning teaching as a career for different types of beginning teachers. *Learning and Instruction*, *18*(5), 408-428.
Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education 63(5),* 376-382.

Acknowledgements

Many thanks to those who have provided encouragement and guidance for this work, including Dan Liston, Jennie Whitcomb, Vicki Hand, Erin Furtak, Joe Polman, Deborah Whitehead, and Nancy Parker.

How Do Learners Process Information in Lectures? The Role of Projected Slides and Type of Note-taking

Christof Wecker, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München, Germany, christof.wecker@psy.lmu.de

Abstract: Today's lectures are typically supported by means of computer-based slide projections, and it is common for learners to take notes on printed handouts containing the lecturer's slides. In a 2x2-factorial design involving 81 students, the effects of slide projection (absent vs. present) and type of note-taking (on empty sheets vs. on handouts) were investigated with respect to the learners' processing of information and their knowledge about information presented on slides and about information presented orally immediately after a lecture and two weeks later after an additional review phase. Preliminary analyses indicate a beneficial effect of note-taking on handouts on knowledge about information presented on slides after the additional review phase, which appears to be mediated to a substantial extent by the processing of information presented on slides while watching the lecture. Hence the effect cannot be explained exclusively by the completeness of this information in the learners' notes.

Today's lectures typically involve computer-based slide projections. In conjunction with this technological aid, students often receive slides in advance to print and use them as a basis for note-taking. Despite substantial literatures about the effects of using computer-based slide projections (see Craig & Amernic, 2006; Levasseur & Sawyer, 2006; Shapiro, Kerssen-Griep, Gayle, & Allen, 2006 for reviews) and the effects of student note-taking in lectures (see Kiewra, 1989; Kobayashi, 2005; 2006 for reviews), little research has focused on the way in which these two aspects interact when using printed slides for note-taking. In particular, the role of the learners' processing of the information presented on slides or only orally in the lecture as well as the role of students' notes for review when preparing for an examination have not been investigated systematically so far. This interplay of learner-generated artefacts and their cognitive processing in a technology-supported learning scenario, both when creating these artefacts and when using them for further study, makes this topic an interesting challenge for Learning Sciences research.

Prior research

Computer-Based Slide Projections

Since the mid-nineties, many researchers have conducted studies about the effectiveness of computer-based slide projections in their own or their colleagues' lectures. According to several reviews, however, this body of research has failed to produce conclusive evidence for beneficial effects of using computer-based slide projections in lectures on student learning (Craig & Amernic, 2006, p. 150; Levasseur & Sawyer, 2006, p. 109-111; Shapiro et al., p. 69; Wecker, 2013). To the contrary, single studies have provided some indication that using computer-based slide projections in lectures may actually be detrimental with respect to certain aspects of knowledge acquisition, i. e. the acquisition of knowledge about information that is presented only orally (Savoy, Proctor & Salvendy, 2009, p. 863; Wecker, 2012, p. 267). Although replication studies are certainly needed to examine the robustness of these negative effects, it may at least be concluded from this research that future work should differentiate knowledge according to the sources of the information that students are supposed to learn about (Wecker, 2012, p. 271).

There are several potential explanations for the negative effect of computer-based slide projections on knowledge about information presented only orally: From a cognitive load perspective, this finding might be explained as a so-called "redundancy effect" (Sweller, 2005) because of cognitive overload caused by simultaneous written and oral presentation of information. On top of several theoretical reasons that question this explanation, the pattern of findings concerning cognitive load did not support it either (Wecker, 2012, p. 268). Another potential explanation for such a negative effect of computer-based slide projections on knowledge about information presented orally is that learners might process information presented only orally less deeply in the presence of computer-based slide projections and thereby less likely comprehend individual sentences (cf. Kintsch, 1988, pp. 166-168; 1998, pp. 96-101). Put differently, learners may be less likely to actually take notice of what is presented to them in oral speech when confronted with computer-based slide projections. Finally, learners may process information presented orally equally well with and without computer-based slide projections, but simply regard the information presented only orally as "secondary" and maybe less important (Adams, 2006). This could be reflected in different amounts of notes taken by the learners about information presented on slides and information presented only orally.

To evaluate these potential explanations, it would be helpful to measure the processing of information presented only orally, the degree of distraction by the processing of information presented on slides, and the amount of note-taking about information presented on slides and presented only orally.

Note-Taking

Potential obstacles to the processing of information presented only orally could be compensated by relieving learners from the requirement to take notes about the information presented on slides. This is what happens when learners receive the lecturers' notes in advance in order to print them and use them as a basis for their note-taking during lectures, as it is common in many universities today.

When considering potential effects of different types of note-taking, the two main functions of note-taking differentiated in the literature need to be taken into account: First, it has been assumed that during note-taking, the processing of the information presented may be beneficially affected due to the requirement to transform it into something that can be written down quickly and still be understood later ("encoding function"). However, research indicates that these mechanisms alone may not be sufficient to improve learning (Kiewra, 1989, pp. 149 f.; Kobayashi, 2005, p. 251).

Second, note-taking enables learners to review and study the information presented to them on a later occasion („storage function"). This characteristic has in fact been found to increase learning outcomes (Kiewra, 1985, p. 33; 1989, p. 148). As a consequence, when studying the effects of different types of note-taking, it is imperative to include measures of learning outcomes after an opportunity for the learners to review their notes.

Several findings are of importance when considering the potential role of note-taking on handouts containing the lecturer's slides: First, the completeness of the learners' notes correlates with learning outcomes (Kiewra, 1987, p. 242; 1989, p. 150). This suggests that note-taking on handouts may foster learning to the extent that it leads to more complete notes.

Second, if learners take notes on their own, i. e. if they do not get handouts but use empty sheets of paper instead, their notes typically cover only a relatively small proportion of the information presented (Kiewra, 1985, p. 33). If, however, information is presented to them in written form (e. g. on the blackboard), it is more likely to be included in the learners' notes than if it is presented to them only orally (Locke, 1977, p. 94). The same has been found for computer-based slide projections (Austin, Lee & Carr, 2004, pp. 317 f.).

Third, if learners receive some kind of support for note-taking, such as incomplete handouts (Cardetti, Khamsemanan & Orgnero, 2010, pp. 82-84) or handouts containing structural outlines (Kiewra, Benton, Kim & Risch, 1995, pp. 175 f.), they acquire more knowledge than if they have to take notes on their own (Cardetti, Khamsemanan & Orgnero, 2010. p. 87; Kiewra, 1985, p. 35; Kiewra, 1989, p. 160; Kiewra, Benton, Kim & Risch, 1995, p. 177). A small number of studies have investigated the effects of note-taking on handouts that contain the lecturer's computer-based slides. Although learners tend to write less in this case than if they have to take notes on their own (Marsh & Sink, 2010, pp. 697; 701), typically the coverage of the information presented is higher (Austin, Lee & Carr, 2004, pp. 317 f.; Stefanou, Hoffman & Vielee, 2008, p. 11). The findings concerning effects on knowledge acquisition remain inconclusive, however (Bowman, 2009, p. 106; Marsh & Sink, 2010, p. 702; Raver & Maydosz, 2010, p. 194).

In light of these considerations it seems appropriate to investigate whether note-taking on empty sheets of paper may be associated with less processing of information presented only orally compared to note-taking on handouts, especially in the presence of a computer-based slide projection that learners attempt to copy down. In contrast, note-taking on handouts might relieve learners of this task and thereby increase the processing of information presented only orally.

Furthermore, note-taking on handouts is likely to lead to higher knowledge acquisition provided that learners have the opportunity to review the material on the basis of the notes they took on the handouts. This should apply particularly to the information presented on slides that is printed on the handouts as well. Such an effect could be due either to the completeness of information presented on slides in learners' notes taken on handouts as compared to less complete notes taken on empty sheets of paper, or to increased processing of information presented on slides while listening to a lecture and taking notes.

Research Questions

The study presented in this paper aims to shed light on how slide projections and the type of note-taking impact learning in lectures via the processing of the information presented on slides and only orally, and via the notes learners take and use for later review. In the following, analyses pertaining to the following three specific research questions will be presented:

(1) What are the main and interaction effects of slide projection and type of note-taking on knowledge about information presented on slides and about information presented only orally immediately after a lecture and after learners had the opportunity to review and study using their own notes?

(2) What are the main and interaction effects of slide projection and type of note-taking on learners' processing of information presented on slides and of information presented only orally during a lecture?

(3) How is the processing of information presented on slides and of information presented only orally during a lecture related to their knowledge about information presented on slides and about information presented only orally immediately after a lecture and after learners had the opportunity to review and study using their own notes?

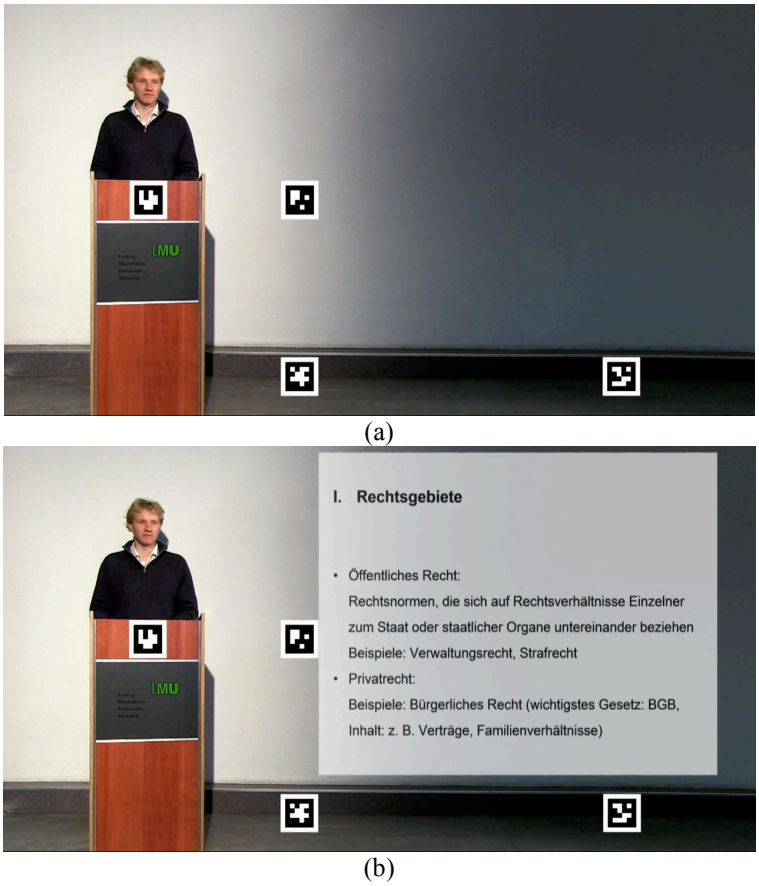
Data analysis for this project has not been finished yet. Therefore, analyses based on about 2/3 of the final sample are included in this paper, while the complete findings will be available for presentation at the conference.

Method

Instructional Unit and Setting

The present study was designed to make learners' processing during lectures accessible to measurement and analysis and to achieve as much experimental control as possible, while at the same time retaining as much of the authentic learning scenario as possible. The approach to measuring learners' cognitive processing of information presented only orally requires interrupting the lecture for the administration of test items referring to information presented within a narrow window of time before the interruption. Experimental control requires minimizing variation in implementation and presentation times among conditions. To achieve both of these goals, the video recording of a lecture was presented to the learners. Retaining as much of the authentic learning scenario as possible implied that learners should not watch the video on a computer screen, but in an almost life-size projection to the wall with high quality.

The topic of the lecture was the German legal system. This topic is covered in many different educational institutions, such as in social studies in ordinary schools or professional schools, in degree programs in Law at the university level, or in adult education centres. Nevertheless, the topic is sufficiently difficult, and the target group of participants has only little knowledge about it. It involves different kinds of knowledge that are typically presented using computer-based slide presentations, such as technical terms, definitions, classifications, rules, conditions for application etc. The lecture was given by a young man standing behind a lectern next to a white area that could be filled by projected slides (see Figure 1a), and had a duration of 30 minutes.



(a)

(b)

Figure 1. Stills from two versions – (a) without versus (b) with projected slides – of the same video recording of a lecture about the German legal system.

To investigate the role of learners' notes for learning in lectures, the participants could study their notes in a second experimental session two weeks after watching the lecture before taking a final test. To this purpose the notes they had taken while watching the lecture were returned to them during this second session for a review phase of 45 minutes. During this phase they could make annotations and take further notes using a pen in a colour different from the one used in first session.

Design

The study had a two-factorial experimental between-subjects design with the factors *slide projection* (absent vs. present) and *type of note-taking* (on empty sheets vs. on handouts). Individual persons were assigned to the four conditions at random.

Procedure

The first session of data collection was conducted with with each learner individually. After a 5-minute introduction, they had up to 15 minutes to complete the first pretest for prior knowledge and further control variables. Then, the gaze-tracking equipment was calibrated, which could last for up to 15 minutes. Next, the learners watched the video recording of the lecture. Depending on whether their processing of the information presented was measured (see below), this could last either 30 minutes or about 45 minutes. Finally, they completed a 40-minute immediate posttest for knowledge and some additional variables.

About two weeks (i. e. 10 to 18 days) later, the learners participated in the second session of data collection in groups of six or seven persons. After completing a second pretest for further control variables within 30 minutes, they reviewed their own notes for 45 minutes. Finally, they completed the 45-minute delayed posttest for knowledge and some additional variables.

Manipulation of Independent Variables

Slide Projection

In the conditions *with slide projection*, text slides were cut into the video recording of the lecture at the position of the white wall area next to the speaker (see Figure 1b). The slides contained technical terms, definitions, rules and principles as well as examples in short phrases rather than full sentences. Overall, 31 slides including a title slide and an outline slide were shown. Each slide comprised at most seven lines (excluding headlines) or six points in bulleted or numbered lists. The background of the slides was white, and the slides did not contain any graphical corporate identity elements.

In the conditions *without slide projection*, the uncut video recording of the lecture was shown to the learners (see Figure 1a).

Type of Note-Taking

In the conditions with *note-taking on empty sheets*, the learners received eight sheets of paper with lines on one side in portrait format. Upon request they could receive additional sheets of the same type.

In the conditions with *note-taking on handouts*, the learners received eight sheets of paper printed with content on the left and lines on the right on one side in portrait format. The content on the left hand side of each sheet corresponded to four slides used in the conditions with projected slides. However, no slide borders were printed on the handouts, and headlines were not repeated (as in sequences of slides belonging to the same section) in order to avoid too close resemblance to slides, especially for the participants in the condition without slide projection and with note-taking on handouts. In addition, upon request the learners could receive sheets of ruled paper of the same type as in the condition with note-taking on empty sheets.

Measurement of Dependent Variables

Knowledge

Knowledge was measured in both the immediate and the delayed posttest by means of an identical computer-based test that contained 32 multiple-choice items with four options one of which was correct. These items covered only content presented during the first two thirds of the lecture, for reasons to be explained below. Half of the items covered information that was presented on slides in the conditions with slide projection, half of them covered information that was presented only orally in all conditions. Separate scales were constructed for *knowledge about information presented on slides* and *knowledge about information presented only orally* by adding the scores for all items belonging to the respective scale. Then, the scales were z-standardized for both the immediate and the delayed posttest using the means and standard deviations from the condition without slide projection and with note-taking on empty sheets in the immediate posttest. Scales for *overall knowledge* in the immediate and delayed posttest were constructed by calculating the mean of the scales for knowledge about information presented on slides and knowledge about information presented only orally in the respective test.

A subset of six items from the posttests was used to measure *prior knowledge* in the first pretest and scored accordingly (but without the z-standardization).

Processing of Information

Processing of information was measured in two ways: Processing of information presented on slides and of information presented only orally was captured by means of items that were interspersed during the final third of the video-recorded lecture (to avoid interference with the knowledge measured in the posttest). In addition, processing of information presented on slides was also captured by means of gaze-tracking.

Processing of both kinds of information was measured by 16 multiple-choice items with four options one of which was correct. These items were presented to the learners during an interruption of the video within 30 seconds after the information covered in the item had been presented in the lecture. Half of them covered information presented on slides, and half of them covered information presented only orally. These constraints were applied both to prevent forgetting and interference of reliance on the phonological loop, and to avoid sensitizing the participants to one of the two sources of information (slides or oral speech). Furthermore, to check for any reactivity associated with this measurement, measures of processing by means of interspersed items were taken from only half of the participants in each condition.

Two scales were constructed for *processing of information presented on slides* and *processing of information presented only orally* by adding the scores for all items belonging to the respective scale. Then, the scales were z-standardized using the means and standard deviations from the participants to whom the interspersed items were administered in the condition without slide projection and with note-taking on empty sheets.

In addition to this item-based measurement of learners' processing of information, their gazes were tracked using an Ergoneers 25 Hz head-mounted eye-tracker. The gaze data were anchored to the visual scenario by means of five marker symbols placed at different positions the learners could look at within the setting, such as the lectern or the projection area in the video, or the learners' table (see Figure 1a and b). The markers were captured by the field camera and recognized automatically by the software.

Areas of interest were defined to capture learners' visual attention to all individual information units on the slides. These areas of interest were anchored to the markers in the field camera video. Furthermore, learners' own note-taking was operationalized as the automatic recognition of the marker placed on the learners' table in the field camera video because this marker became only visible when the learners lowered their eyes to their notes. Besides, for each information unit in the lecture the time windows in which they were presented orally to each participant were determined based on timestamps in the log files.

These data are used to construct the following measures: (a) the proportion of each learner's dwell time in the corresponding area of interest during which the same information unit is presented orally (indicating the *proportion of slide reading with consistent oral input* for each information unit), (b) the proportion of each learner's dwell time in the corresponding area of interest during which a different information unit is presented orally (indicating the *proportion of slide reading with inconsistent oral input*), (c) the proportion of the oral presentation time during which the learner read the text in the corresponding area of interest (indicating the *proportion of oral input consistent with slide reading*), and (d) the proportion of the oral presentation time during which the learner read the text in a different area of interest or was occupied with his or her notes (indicating the *proportion of oral input inconsistent with slide reading*). These analyses will be completed by the time of the conference.

Amount of Note-Taking

The learners' notes are used as a further data source. The amount of note-taking is determined separately for information that is presented on slides and information presented only orally. The learners' notes are segmented to idea units of a size (typically sentences or phrases) corresponding to the 261 information units contained in the lecture. Each of the resulting segments is coded as either one of the 261 information units contained in the lecture or as other content. Furthermore, each segment is coded as written down while watching the lecture or as written down during the review phase, based on the colour of the pen used.

Three coders independently analyzed a portion of the data from all conditions. Their agreement indicates satisfactory objectivity (81 to 82 %; Cohens $\kappa = .81$ to $.82$). The whole material is currently being coded.

Two scales are constructed from these data: The *amount of note-taking concerning information that is presented on slides* and the *amount of note-taking concerning information that is presented only orally* are calculated as the proportion of the information units contained in the learner's notes as compared to all information units from the respective information source (slides or only oral speech).

Participants

The dataset analyzed for the present study includes the first 81 participants, mainly students of Psychology or Education (but not of Law). They were recruited by means of lecture visits, postings on notice boards, and social media. They were compensated by receiving either money or credit counting towards participation requirements in their degree programs.

The sample comprises 73 % female and 27 % male students. On average they were $M = 24.0$ ($SD = 6.1$) years old.

Statistical Analysis

Because of the preliminary character of the present analyses based on a sample smaller than the final sample size, the level of significance was set to 10 % for all statistical tests presented in this paper.

Results

Effects on Knowledge about Information Presented on Slides and Knowledge Presented Only Orally Immediately after the Lecture and Two Weeks Later

The effects of slide projection and type of note-taking on knowledge about information presented on slides and knowledge presented only orally immediately after the lecture and two weeks later were analyzed by means of analyses of covariance with prior knowledge as a covariate. In addition to the two independent variables, the factor whether measurement of the processing of information by means of interspersed items had taken place was included as a further between-subjects factor. With one exception, no significant main effect of the measurement of the processing of information and no interaction effect with any of the other factors was found.

Concerning *knowledge about information presented on slides in the immediate posttest*, learners in the conditions without slide projection ($M = 0.10$; $SD = 0.82$) demonstrated only minimally higher amounts of knowledge about information presented on slides than learners in the condition with slide projection ($M = 0.07$; $SD = 0.69$). This difference was not significant, $F(1; 72) = 0.52$; $p = .48$; partial $\eta^2 = .01$. Similarly, learners in the conditions with note-taking on empty sheets ($M = 0.12$; $SD = 0.80$) exhibited little more of this knowledge than learners in the conditions with note-taking on handouts ($M = 0.06$; $SD = 0.70$). This difference was not significant, $F(1; 72) = .34$; $p = .56$; partial $\eta^2 = .01$, and neither was the interaction of the two independent variables, $F(1; 72) = 1.48$; $p = .23$; partial $\eta^2 = .02$.

Although with respect to *knowledge about information presented only orally in the immediate posttest* the descriptive findings showed slightly higher knowledge about information presented only orally in the conditions without slide projection ($M = 0.19$; $SD = 0.93$) as compared to the conditions with slide projection ($M = 0.01$; $SD = 1.07$) – which is in line with the speech suppression effect observed in the literature –, this difference was not significant, $F(1; 72) = 1.91$; $p = .17$; partial $\eta^2 = .03$. Likewise, the difference between the conditions with note-taking on empty sheets ($M = -0.10$; $SD = 1.05$) and the conditions with note-taking on handouts ($M = 0.28$; $SD = 0.94$) was not significant, $F(1; 72) = 1.93$; $p = .17$; partial $\eta^2 = .03$. Furthermore, there was no indication of an interaction effect, $F(1; 72) = 0.05$; $p = .83$; partial $\eta^2 < .01$

Knowledge about information presented on slides in the delayed posttest was minimally lower in the conditions without slide projection ($M = 0.65$; $SD = 0.83$) than in the conditions with slide projection ($M = 0.69$; $SD = 0.87$), but this difference was not significant, $F(1; 72) = 0.02$; $p = .88$; partial $\eta^2 < .01$. In contrast, the difference between the conditions with note-taking on empty sheets ($M = 0.36$; $SD = 0.82$) and note-taking on handouts ($M = 0.95$; $SD = 0.78$) in favour of the latter was significant and corresponded to a large effect, $F(1; 72) = 10.60$; $p < .01$; partial $\eta^2 = .13$. The interaction effect was not significant, $F(1; 72) = 2.05$; $p = .16$; partial $\eta^2 = .03$.

With respect to *knowledge about information presented only orally in the delayed posttest* learners in the conditions without slide projection ($M = 0.45$; $SD = 0.97$) demonstrated slightly higher knowledge about information presented only orally than learners in the conditions with slide projection ($M = 0.31$; $SD = 0.98$). Although again descriptively in line with the speech suppression effect, this difference failed to reach significance, $F(1; 72) = 1.37$; $p = .25$; partial $\eta^2 = .02$. The difference between the conditions with note-taking on empty sheets ($M = 0.24$; $SD = 1.03$) and note-taking on handouts ($M = 0.50$; $SD = 0.90$) in favour of the latter was not significant as well, $F(1; 72) = 0.82$; $p = .37$; partial $\eta^2 = .01$. Again, there was no indication of an interaction effect, $F(1; 72) = 0.41$; $p = .52$; partial $\eta^2 = .01$.

Effects on Processing of Information Presented on Slides and Information Presented Only Orally

The effects of slide projection and type of note-taking on the processing of information presented on slides and knowledge presented only orally were analyzed by means of analyses of covariance with prior knowledge as a covariate on the basis of the learners that received the interspersed items. The *processing of information*

presented on slides was slightly higher in the conditions without slide projection ($M = 0.31$; $SD = 1.49$) than in the conditions with slide projection ($M = 0.11$; $SD = 1.65$), but this difference was not significant, $F(1; 33) = 0.41$; $p = .53$; partial $\eta^2 = .01$. However, learners in the condition with note-taking on empty sheets ($M = -0.34$; $SD = 1.60$) – according to the less stringent level of significance applied in this preliminary analysis – exhibited significantly lower processing of information presented on slides than learners in the condition with note-taking on handouts ($M = 0.63$; $SD = 1.44$), $F(1; 33) = 3.12$; $p = .09$; partial $\eta^2 = .09$. No significant interaction effect was found, $F(1; 33) = 0.52$; $p = .48$; partial $\eta^2 = .02$.

With respect to the *processing of information presented only orally* the conditions without slide projection ($M = 0.05$; $SD = 1.24$) and with slide projection ($M = -0.08$; $SD = 1.53$) did not differ significantly, $F(1; 33) = 0.32$; $p = .57$; partial $\eta^2 = .01$. Also the difference between the conditions with note-taking on empty sheets ($M = -0.36$; $SD = 1.19$) and the conditions with note-taking on handouts ($M = 0.25$; $SD = 1.53$) did not differ significantly in this respect, $F(1; 33) = 1.23$; $p = .28$; partial $\eta^2 = .04$. Furthermore, there was no significant interaction effect, $F(1; 33) = 0.89$; $p = .35$; partial $\eta^2 = .03$.

The Role of Processing of Information for Knowledge

The processing of information presented on slides significantly correlated with knowledge about information presented on slides both in the immediate posttest, $r = .35$; $p = .02$, and in the delayed posttest, $r = .59$; $p < .01$. Similarly, the processing of information presented only orally significantly correlated with knowledge about information presented only orally both in the immediate posttest, $r = .52$; $p < .01$, and in the delayed posttest, $r = .55$; $p < .01$.

As the type of note-taking had a significant effect on knowledge about information presented on slides in the delayed posttest, a mediation analysis for this effect was conducted with processing of information presented on slides as the potential mediator. The effect size of partial $\eta^2 = .21$ without controlling for processing of information presented on slides that was found for the learners from whom measures of processing were collected was reduced to partial $\eta^2 = .14$ by including processing of information presented on slides as a covariate. This means that 32.9 % of the effect of note-taking on handouts as compared to note-taking on empty sheets on knowledge about information presented on slides exhibited in the delayed posttest seem to be mediated by the processing of information presented on slides while watching the lecture.

Discussion

The descriptive differences between the conditions without and with slide projection with respect to knowledge about information presented only orally in both posttests are compatible with the speech-suppression effect reported in the literature (Wecker, 2012, p. 267). However, in the preliminary analyses presented here they failed to reach significance. It remains to be seen whether in the final sample this effect will be replicated.

Note-taking on handouts had a strong beneficial effect compared to note-taking on empty sheets on knowledge about information presented on slides (and printed on the handouts) after the learners had the opportunity to study using their notes, and tended to benefit the processing of information presented on slides during the lecture. The former effect appears to be mediated to a large extent by the processing of information presented on slides while watching the lecture. This suggests that – in contrast to the storage hypothesis – this effect is not exclusively due to the availability of the information that was presented on slides on the handouts during review. Rather – in line with the encoding hypothesis – learning seems to be influenced by the processing of this information during note-taking as well (Kiewra, 1989, pp. 149 f).

That the two scales for processing of information presented on slides and of information presented only orally significantly correlated with the respective knowledge scales in the immediate and in the delayed posttest, provides some initial validation for these measures.

Obvious limitations of the analyses presented here include the liberal level of significance chosen due to the current state of the project. This issue is connected to the insufficient sample size included in the present analyses as well as the issue of not yet including measures of processing of information based on gaze tracking and analyses of the amount of note-taking. Beyond these limitations that have to do with the current state of the project, despite serious efforts to keep the learning situation as similar to a real-life lecture setting, several characteristics of the learning scenario may be regarded as somewhat artificial. In particular, this applies to the video-taped lecture viewed in individual learning sessions and interrupted for some participants by interspersed items, to the review phase of limited duration immediately before the final test, during which the learners could only use their notes rather than textbooks and other resources, and the final test as a proxy for a real examination. Future studies should relax some of the laboratory-type restrictions of the present study.

References

Adams, C. (2006). PowerPoint, habits of mind, and classroom culture. *Journal of Curriculum Studies, 38*(4), 389-411.

Austin, J. L., Lee, M. & Carr, J. P. (2004). The effects of guided notes on undergraduate recording of lecture content. *Journal of Instructional Psychology, 31*(4), 314-320.
Bowman, L. L. (2009). Does posting PowerPoint presentations on WebCT affect class performance or attendance? *Journal of Instructional Psychology, 36*(2), 104-107.
Cardetti, F., Khamsemanan, N. & Orgnero, M. C. (2010). Insights regarding the usefulness of partial notes in mathematics courses. *Journal of the Scholarship of Teaching and Learning, 10*(1), 80-92.
Craig, R. J. & Amernic, J. H. (2006). PowerPoint presentation technology and the dynamics of teaching. *Innovative Higher Education, 31*, 147-160.
Kiewra, K. A. (1985). Providing the instructor's notes: An effective addition to student notetaking. *Educational Psychologist, 20*(1), 33-39.
Kiewra, K. A. (1987). Notetaking and review: the research and its implications. *Instructional Science, 16*, 233-249.
Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*, 147-172.
Kiewra, K. A., Benton, S. L., Kim, S. I. & Risch, N. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology, 20*, 172-187.
Kintsch, W. (1988). The role of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*(2), 163-182.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*(2), 242-262.
Kobayashi, K. (2006). Combined effects of note-taking/reviewing on learning and the enhancement through interventions: A meta-analytic review. *Educational Psychology, 26*(3), 459-477.
Levasseur, D. G. & Sawyer, J. K. (2006). Pedagogy meets PowerPoint: a research review of the effects of computer-generated slides in the classroom. *Review of Communication, 6*(1), 101-123.
Locke, E. A. (1977). An empirical study of lecture notetaking among college students. *The Journal of Educational Research*, 71(2), 93-99.
Marsh, E. J. & Sink, H. E. (2010). Access to handouts of presentation slides during lecture: Consequences for learning. *Applied Cognitive Psychology, 24*, 691-706.
Raver, S. A. & Maydosz, A. S. (2010). Impact of the provision and timing of instructor-provided notes on university students' learning. *Active Learning in Higher Education, 11*, 189-200.
Savoy, A., Proctor, R. W. & Salvendy, G. (2009). Information retention from PowerPoint and traditional lectures. *Computers & Education, 52*(4), 858-867.
Shapiro, E., Kerssen-Griep, J., Gayle, B. & Allen, M. (2006). How powerful is PowerPoint? Analyzing the educational effects of desktop presentational programs in the classroom. In B. M. Gayle, R. W. Preiss, N. Burrell & M. Allen (Eds.), *Classroom communication and instructional processes: Advances through meta-analysis* (pp. 61-75). Mahwah, NJ: Erlbaum.
Stefanou, C., Hoffman, L. & Vielee, N. (2008). Note-taking in the college classroom as evidence of generative learning. *Learning Environments Research, 11*, 1-17.
Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 151-167). Cambridge: Cambridge University Press.
Wecker, C. (2012). Slide presentations as speech suppressors: When and why learners miss oral information. *Computers & Education, 59*(2), 260-273.
Wecker, C. (2013, August). *Does PowerPoint foster learning? A meta-analysis*. Paper presented at the 15th Biennial Conference for Learning and Instruction of the European Association for Research on Learning and Instruction (EARLI) "Responsible Teaching and Sustainable Learning", Munich, 27–31 August 2013.

Acknowledgments

This research has been funded by a grant (WE 5426/1-1) from the Deutsche Forschungsgemeinschaft (DFG).

Using Analytics for Improving Implementation Fidelity in an Large Scale Efficacy Trial

Mingyu Feng, Jeremy Roschelle, Robert Murphy
SRI International, 333 Ravenswood Ave, Menlo Park CA 94025, USA
Email: mingyu.feng@sri.com, jeremy.roschelle@sri.com, robert.murphy@sri.com

Neil T. Heffernan, Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609
nth@wpi.edu

Abstract: The field of learning analytics is rapidly developing techniques for using data captured during online learning. In this article, we develop an additional application: the use of analytics for improving implementation fidelity in a randomized controlled efficacy trial. In an efficacy trial, the goal is to determine whether an innovation has a beneficial effect under best-case implementations. Analytics is more accurate and less expensive than traditional ways of collecting and analyzing implementation fidelity data, and may allow targeted adaptations of the innovation that improve the quality of the research. We report our experience in developing and using analytics during the course of an efficacy trial that evaluated the use of ASSISTments as an online homework tool for middle school mathematics.

Significance

The fields of learning analytics and educational data mining are rapidly developing new techniques for using the copious data that is captured during online learning. Applications of learning analytics have included predicting student outcomes, improving learning resources, and intervening for particular students to enhance their learning trajectories. In this article, we develop an additional application: the use of analytics as a technique for improving implementation fidelity in a randomized controlled efficacy trial.

In an efficacy trial, the goal is to determine whether an innovation has a beneficial effect under best-case implementations. An important contrast is to an effectiveness trial, which aims to measure effects when the innovation is in broader use, with more environmental variation and less control over implementation. Because of the emphasis on best-case interventions in an efficacy trial, it is fair game in an efficacy trial to monitor and adjust implementation of the innovation. Analytics, we will argue, can provide an important new tool for monitoring implementation fidelity, and thus can allow targeted adaptations of the innovation that improve the quality of the research.

Conducting a Randomized Controlled Trial (RCT) is the primary methodology for educational efficacy research. In its basic form, an RCT randomly assigns participants to alternative conditions, where the conditions are deliberately designed to emphasize a desired contrast. Outcomes are measured, and if there is a difference in outcomes in the contrasting conditions, then the contrasting features of the two conditions are presumed to cause the difference. This inferential process depends on the quality of contrast as experienced by the participants: if the contrast is weak, or highly variable, or drifts away from the intended contrast, then measured effects may be due to something other than the designed contrast. Thus, it is important to understand the contrast between conditions as implemented, which traditionally leads to the idea of implementation fidelity—are the conditions implemented in a way that highlights the planned difference between conditions? Is the treatment condition being implemented in a way that preserves the potential for a beneficial effect?

When an efficacy trial is conducted in schools, collecting and analyzing implementation fidelity data is typically slow, inaccurate, and/or expensive. Indeed, often the analysis of implementation fidelity only occurs after the experiment is complete—which can be wasteful if it turns out that the desired contrast was not implemented well, and therefore the experiment is invalid (i.e. the investigators can obtain "no effect" because the treatment was not implemented well according to the model specified by innovation developers, not because the treatment could not produce benefits). Traditional methods for collecting implementation fidelity data are through observations or through self-report. Observations are expensive to conduct, and in contexts where 50 or 100 schools participate in a study, which is usually the case for an efficacy trial, it is typical that projects can only afford one or two observations per school year. Self-report is less expensive to collect, but can be inaccurate. Retrospective interviews are a third source of data, but also introduce concerns about inaccuracy or biases. We argue that analytics are an additional method for collecting implementation fidelity data that can be faster, cheaper, and more objective. In support of this claim, we report our experience in developing and using analytics during the course of an RCT that evaluated the use of ASSISTments as an online homework tool for middle school mathematics in the state of Maine.

 © ISLS

Methodological Approach

Implementation Fidelity

Implementation Fidelity is the extent to which the delivery of an intervention conforms to the protocol or program model as intended by the developers of the intervention (Domitrovich & Greenberg, 2000; Mowbray et al., 2003). The assessment of implementation fidelity has been highlighted as critical to understanding how programs are implemented in efficacy studies (Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008; Dusenbury et al., 2003). Despite the value of measuring implementation fidelity in conducting and interpreting RCTs, a long-standing problem is that implementation fidelity is often not measured or underreported, likely due to the expense and difficulty of collecting relevant data.

There have also been critiques of the construct of implementation fidelity: it seems to assume that an innovation should be delivered in the same way in every school, and may better suit over-scripted approaches than highly adaptive approaches. What if implementing a particular innovation "well" means engaging in extensive adaptations of the innovation? As our case study of ASSISTments will show, this concern can be addressed: ASSISTments is intended to be highly adaptive and yet it still makes sense to monitor fidelity, for example, by monitoring whether teachers are using the ASSISTments facility for adapting homework problem sets. Thus it seems possible to develop analytics that detect adaptive or non-adaptive behavior by teachers, and such detectors can contribute to understanding of whether the expected adaptations are likely to be occurring with the innovation.

The literature favors *model-based* approaches, which consider implementation fidelity relative to a logic model (Nelson et al. 2010). A typical logic model traces the causal pathway(s) from affordances designed on the basis of learning theory, to inputs provided to a school (such as new software or teacher professional development), to activities enacted in the school using the inputs, to outcomes that are measured. A sound logic model is central to any high quality efficacy trial.

When measuring implementation fidelity relative to a model, five types of implementation information may be helpful (Cordray, 2008): adherence, exposure, quality of delivery, participant responsiveness, and program differentiation (Durlak & DuPre, 2008; Dusenbury et al., 2003; Fagan et al., 2008). A first pair of implementation measures addresses availability and use of inputs: *Adherence* tracks whether the expected inputs are actually in use at the target schools: do participants access and use the resources provided? *Exposure* monitors how much of the resource is used: is the full extent of the resource used? Are the frequency and dosage of use as intense as the developer recommends? A second pair of implementation metrics addresses the quality of the ensuing activities at schools. *Quality of delivery* reflects the manner in which a program is delivered and can capture whether the activities using the resources are unfolding according to the expected teaching and learning processes. For example, if software attempts to give students practice using the "spacing effect" as recommended by Pashler et al. (2007), are students actually practicing the same skills at regularly spaced intervals? *Participant responsiveness* can look at uptake by teachers and students of the features of the innovation: for example, if the system provides teachers with reports, do they open them? If students have opportunities to choose more challenging problems or to watch tutorial videos, do they do this? Finally, a last category concerns the intended contrast. *Program differentiation* looks at whether the treatment conditions are different from other conditions in expected ways, including mediating processes. For example, if an innovation is expected to increase overall learning by providing more feedback to learners, are we sure that learners in the control condition are not getting the same levels of feedback, but through different processes? We argue that analytics could be developed for these categories of implementation fidelity.

Learning Analytics

The field of educational data mining and learning analytics (LA) has developed rapidly recently (Baker & Yacef, 2009; Romero & Ventura, 2007, 2010; Siemens & Baker, 2012; U.S. Department of Education, 2012). The 2013 Horizon Report (EDUCAUSE, 2013a) describes learning analytics as the "…field associated with deciphering trends and patterns from educational big data, or huge sets of student-related data, to further the advancement of a personalized, supportive system of higher education." However, LA is not limited to higher education. With technology usage become popular and more accessible among younger children, there has been growing use of LA in K-12 settings (EDUCAUSE, 2013b). The main purpose of LA has been to observe and understand learning behaviors in order to enable appropriate interventions at the individual, course, department, or even institution level (Brown, 2011).

Online learning systems —learning management systems, learning platforms, and learning software— have the ability to capture streams of fine-grained learner behaviors. Then it is the responsibility of data analysts to operate on the data, through procedures such as raw data processing, data aggregation, and/or data modeling using data mining algorithms, in order to make necessary inferences. Different from pure data mining, the process of LA often draws on a broader array of academic disciplines, incorporating concepts and techniques from information science and sociology, in addition to computer science, statistics, psychology, and learning

sciences. A good understanding of the entire learning system and educational environment where the system was used is also needed to draw useful and valid conclusions. Once the data analysis is completed, the findings are provided to a variety of stakeholders who can use the feedback to improve instruction, or improve the learning systems, or other educational decision-making for learners. Thus, the feedback loop is closed.

So far, improving student's learning outcomes has been the core goal of LA. While in this paper, the use of LA supports the closure of a different feedback loop that involves innovation developers, evaluators, and implementers, and implementation supporters.

Case Study: The ASSISTments Efficacy Trial

ASSISTments System and Research Design

ASSISTments (www.assistments.org) is an online tutoring system that provides "formative assessments that assist." Teachers choose (or add) homework items in ASSISTments and students can complete their homework items online. As students do homework in ASSISTments, they receive feedback on the correctness of their answers. Some students may choose to do their homework offline, but in typical use, teachers still require students to upload their answers before coming to class. Some problem types also provide hints on how to improve their answers, or help decompose multistep problems into parts (see Figure 1). Teachers may choose to assign problem sets called "skill builders" that are organized to promote mastery learning (Anderson, 2000). Teachers also receive reports on their students' homework and can use this information to organized more targeted homework reviews, to assign specific follow-up work to particular students, and to more generally adapt or differentiate their teaching. ASSISTments is provided to schools as a free service of Worcester Polytechnic Institute (WPI). Prior research has found that analytics based on students' usage of the system during the year can predict end-of-year scores on statewide standardized test (Feng et al. 2009; Pardos et al. 2013), identify students engagement states (San Pedro et al., 2013) and college attendance (San Pedro et al., 2013b)..

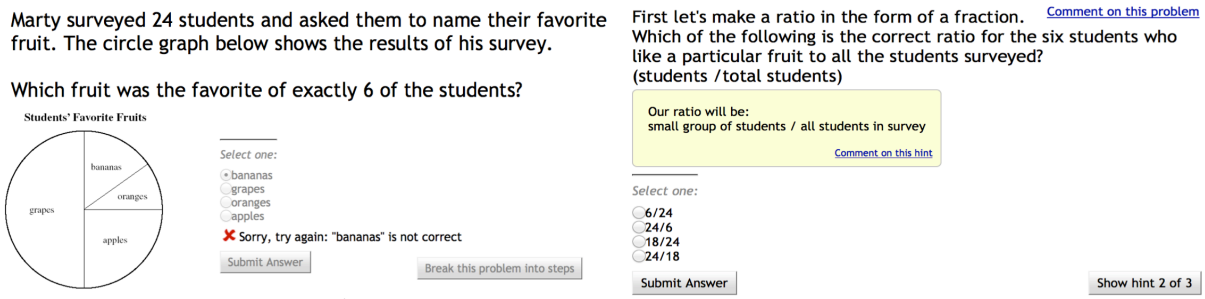


Figure 1. Screen shots of an 7th grade item in ASSISTments that provides correctness feedback and breaks the problem into steps (left) and the first sub-step with a hint message (right)

Prior research also has established the promise of ASSISTments for improving student outcomes in middle school mathematics through homework support (Mendicino, Razzaq, and Heffernan, 2009; Singh et. al, 2011). Building on this prior research, a team led by SRI International in collaboration with WPI and the University of Maine is conducting a large-scale efficacy trial with ASSISTments in the state of Maine where a one-to-one laptop program was well established. The research is an RCT involving 45 middle school schools from two cohorts, with schools randomly assigned to treatment or control (i.e. "business as usual") conditions. The intervention is implemented in Grade 7 math classrooms in treatment schools over 2 consecutive years (academic years 2012–13 and 2013–14 for Cohort 1 schools and 2013–14 and 2014–15 for Cohort 2 schools). In the Treatment condition, teachers receive professional development (PD) and use ASSISTments in the first year to become proficient with the system and then teachers use ASSISTments with a new cohort of students in the second year when student outcomes are measured. Note that we are testing students in *teachers'* second year of experience with the system, because of the developer's belief that teachers do not sufficiently master the system in their first year of experience.

This design provides a strong opportunity for using analytics for implementation fidelity. Since the goal in the first year is to achieve teacher proficiency with the system, if analytics can reveal whether or not this is occurring in a timely manner, additional mentoring could be provided to bring all teachers up to desired levels of implementation. This can occur *before* the second, measurement year begins.

ASSISTments Logic Model

The efficacy trial is guided by the ASSISTments logic model (See Figure 2). Note that the logic model allows for three pathways to increased student learning. A first path is that students may complete homework

with greater regularity when it is online. Even if there was nothing different about homework online or offline, completing more homework could improve student learning. A second path, labeled "direct effect" is the effect on students of getting support for doing homework. A third path is through reporting to teachers, who can then adapt instruction to their students' needs. Our strategy was to align potential implementation fidelity analytics to the "features" and "mediating variables" columns of the logic model.

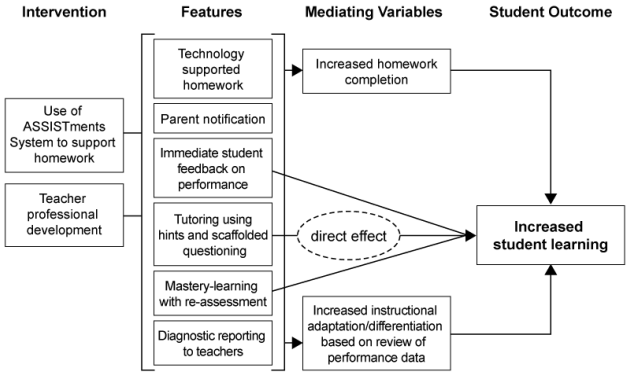


Figure 2. The ASSISTments logic model

Specified Use Model

Based on developers' prior experience in school implementation, we set the specified use model such that teachers who use ASSISTments in the study are expected to assign approximately 25 minutes of homework in ASSISTments for a minimum of 3 nights per week. Homework assignments created by teachers within ASSISTments are expected to consist of: (1) mastery learning problem sets (aka. "skill builders") that addresses a prerequisite or recently-instructed mathematics skill; (2) reassessment mastery problems that are automatically assigned by the system and address a skill that a student has previously mastered; (4) and a series of textbook problems that will comprise the majority of the assignment.

Teachers will receive a performance report early the next morning via email, in addition to other reports that they can access after logging into their ASSISTments account. The report informs teachers whether a student completed the assignment, student's performance on each problem/skill, and also identifies the problems/skills with which most students struggled. Teachers are expected to review ("open") the homework performance report for a minimum of 50% of assignments.

Design of Candidate Analytics

Our data analytics for the first implementation year, as reported in this paper, center on guiding PD and mentoring offered to the teachers in the first year Later on, when data has been collected for the overall RCT, the same analytics may be useful as moderating or mediating variables in the analysis of the impact.

The design of candidate analytics was guided both by the categories of implementation fidelity (e.g. adherence, exposure, quality of delivery, uptake) and by the pathways in the logic model. Below we describe how we used this guidance to design and try a wide variety of analytics.

1. *Adherence*. We were able to determine whether teachers were using the system to assign homework to their students, and whether they were appropriately using homework problems from their textbooks as well as "skill builder" problem sets. We could also see whether students were using the system to access and do homework.
2. *Exposure*. We could see how much homework was being assigned and how often it was assigned. We could see whether students were getting opportunities to use all ASSISTments features or just a subset of features. Another very interesting variable was the time of day when students were using ASSISTments: were they doing homework at home or in school as well?
3. *Quality of Delivery and Uptake*. For teachers, we could detect the "adaptive teaching" route of the logic model by seeing whether teachers were opening reports on their students' homework (as opening the reports is a necessary precursor to adapting instruction on the basis of the reports). We also could detect student uptake and use of the system: how many minutes per week were students using the system? Was this consistent among students with a given teacher, or was it highly variable?

One important limitation of our plans is that teachers in the control schools are not using ASSISTments, and therefore we cannot get comparable data in the control schools. Because of this, the study is still using self-report, interview, and observational measures to get information about control schools and to understand the enacted contrast.

Data Sources and Analysis

In the 2012-2013 school year, 17 schools in the state of Maine were recruited as the first cohort of participants and 9 schools and all 7th grade teachers in these schools were randomly assigned to the treatment condition. Overall, 13 treatment teachers and over 800 7th grade students from their classes used ASSISTments to do their homework. Each teacher and student has his/her own login account. Thus, all actions made in the system can be tracked individually. As a student works online in ASSISTments, the system keeps a detailed log (aka "the click stream") of his/her interaction with the tutor, including answers given, whether correct or incorrect, requests for hint messages, or other interface selections such as clicking on specific links to start an assignment or moving ahead to the next problem. Additionally, an offline version enables students to use ASSISTments even when they don't have Internet access at home. Student work is recorded on their laptops and uploaded to the ASSISTments' server when the laptop is connected to the Internet at school. The offline use data will be included in the reports when the teacher opens then online the next time. Teacher's use of the system such as assigning homework, the type of the assignment, clicking a link to open a specific report, is also logged. All of the actions are time-stamped. To compute the candidate analytics, we collected ASSISTments system log data for the 13 teachers and their students from the period from February to April 2013. The log is fine-grained behavioral and outcome data as students interact with the system.

Measures of treatment fidelity were developed based on the log data to assess the extent to which teachers and students in the treatment condition followed the specified use model, as described above. Our approach to data analysis was essentially descriptive, using aggregated statistical metrics. At this stage of the efficacy trial, a goal was for the data analysis team to present a portrait of implementation to the development team, and to ask: *is this quality of implementation you were expecting to see and would be happy to have tested in with new cohort of student next year? If not, are there actions you can take that might bring implementation up to your desired levels before the next school year starts*?

Findings

A first useful analytic was how often teachers made assignments with ASSISTments. We found that across 3 months, teachers were assigning approximately 1-2 homework assignments per week in ASSISTments (Figure 3). This was lower than the 3 assignments per week that were originally expected. The team also looked at homework completion rates, which were around 75% and average minutes spent doing homework that was round 15 minutes. Both of these values were approximately as expected. Overall, the team felt the rate of homework assignment was a little low, but acceptable given the minutes spent doing homework and the completion rates.

We then looked at the type of assignment (See Figure 4). This revealed that teachers were assigning standard textbook homework problems about half the time and mastery "skill builders" about a quarter of the time. About 25% of the assignments were not from textbook. This was viewed as very promising, as it countered an earlier fear that teachers were sticking with traditional homework items and not using the potentially more powerful mastery problem sets in ASSISTments. We also learned from this analysis how useful analytic trend information can be. The teacher professional development logic for ASSISTments assumed that teachers would start with the more familiar "textbook homework" and gradually feel comfortable to include less familiar "skill builder" problem sets. The trend towards more "skill builder" usage suggested that this was indeed occurring.

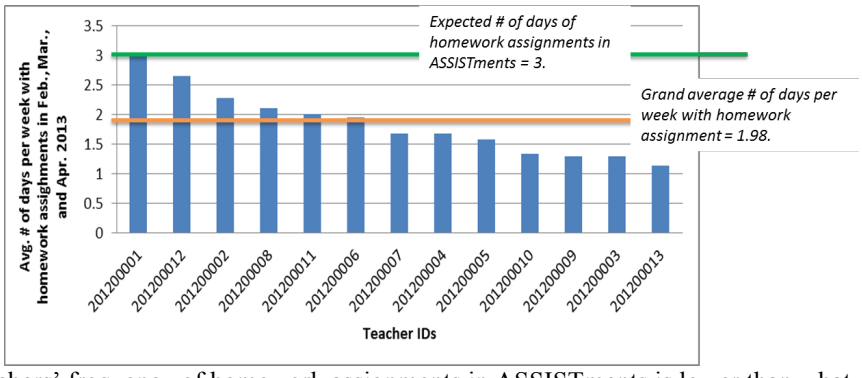


Figure 3. Teachers' frequency of homework assignments in ASSISTments is lower than what was originally expected

		# of Textbook Problem Set assigned			Textbook Problem Set %	# of ASSISTments SkillBuilder assigned			SkillBuilder %	Total # of assignments
School	Teacher ID	Feb	Mar	Apr		Feb	Mar	Apr		
ck Junior High School	201200006	12	18	6	69.23%	3	13		30.77%	52
ck Junior High School	201200009	9	6		34.09%	22	3	4	65.91%	44
nington Elementary School	201200007	1	2	13	33.33%	12	6	14	66.67%	48
nt Elementary School	201200001	26	42	38	100.00%					106
ion Middle School	201200010	4	5	2	78.57%			3	21.43%	14
ion Middle School	201200004	4	4	3	40.74%	4	4	8	59.26%	27
District Middle School	201200012	15	16	10	59.42%	13	8	7	40.58%	69
Valley Middle School	201200003	10	13	2	100.00%					25
iro Community School	201200011	5	6	1	75.00%		4		25.00%	16
iro Community School	201200002	11	12	7	87.50%	1	3		12.50%	34
iro Community School	201200013	3	4	1	57.14%	2	2	2	42.86%	14
im S. Cohen School	201200005	7	5	9	100.00%					21
im S. Cohen School	201200008	11	11	4	66.67%	10	3		33.33%	39
	Grand Total				71.88%				28.13%	544

Figure 4. The type of assignments that teachers have made in ASSISTments

The graph in Figure 5 shows what percentages of problems were solved (on average) in each hour of the day. It is obvious that most use happens during the school time from 8am to 2pm, with some usage in after school hours from 3pm to 9pm. This was unexpected as the system was intended for homework analysis. Through complementary methods (such as teacher interviews), we are seeking to determine why most usage was during school hours.

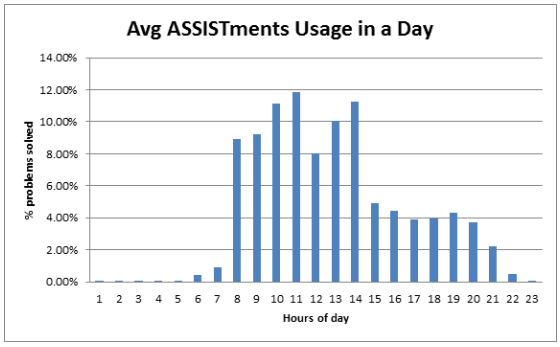


Figure 5. Distribution of student's usage of ASSISTments during a day

A key "uptake" analytic was whether teachers were opening ASSISTments reports, as this is a necessary prelude to adaptive teaching. Here, variation was profound and surprising. The key ASSISTments trainer was very surprised at the particular teachers who were not opening reports; apparently these teachers gave the impression that they were implementing adaptive teaching. We also saw variation within schools; different teachers in the same school were not using the reports equally often. This was confirmed by the field observations that are currently being conducted in schools. These data points led to concrete plans (such as a targeted discussion in an upcoming webinar, or a class visit) to follow up with the teachers who were not yet using the reports often, so as to activate the adaptive teaching pathway in the logic model for all classrooms in the treatment condition.

Conclusion

The quality of efficacy studies can be improved if the expected contrast between conditions can be actively maintained. This is typically difficult in large school-based studies, because of the difficulties associated with self-report (inaccurate), interviews (time consuming and unreliable), and observations (expensive). We explored the utility of creating analytics based on the automatically collected data in the ASSISTments system to address implementation fidelity.

We were able to map each of the three main pathways of the logic model to at least one analytic measure. The adaptive teaching pathway could be examined by looking at whether teachers open reports. The homework completion pathway could be examined by looking at homework completion rates. And the pathway of direct student impacts from having more support while doing homework could be examined by looking at the frequency of homework assignments, the types of problems included in the assignments, and the minutes students were spending doing assignments online. On a whole, this is an extraordinary amount of useful implementation data that we were able to gather and analyze at very modest cost and with high objectivity (especially compared to the inaccuracies of self-report and interview measures).

Further, we were able to design analytics corresponding to four of the five categories. We could look at adherence and exposure by whether teachers were using ASSISTments, how much homework was being assigned, and how many minutes of student usage. We could look at quality of implementation and uptake by seeing which types of problems teachers were assigning and whether they were opening reports. Trend data was particularly useful in showing that quality of implementation was improving over time, which was expected. The major limitation was that we did not have access to comparable data from control conditions, and thus the other more typical ways of collecting implementation data are still necessary.

ASSISTments team members who were coaching with teachers were able to target particular teachers and particular behaviors for their further coaching. Their surprise at which teachers needed coaching indicated the value of combining their own impressions with more objective analytic data. The ASSISTments team also learned from the time-of-day data that students were doing homework not just at night, but also during the school day and in the afternoon.. Further, schools could be encouraged to set up school library computers for students to do homework during the day, which might further increase minutes spent on the system and completion rates.

In the future, we will be able to examine these analytics as mediating variables in our models of outcomes in the RCT. In the past, analytics have been predictive of outcome data. If the study finds that ASSISTments is efficacious, this could be very useful for making recommendations to schools and teachers for further implementation—we may learn that certain characteristics of usage best predict outcomes (such as number of minutes used) and this could guide schools in their further implementations. Considerable work remains to be done to more thoroughly validate particular analytic approaches; for example, it would be useful to compare interview or observational data to system-based measures. It could be, for example, that some teachers are assigning a mix of homework both within and outside ASSISTments, which could lead to different interpretations of how to intervene to increase implementation fidelity. Yet even at the descriptive level addressed here, the analytic data was perceived as very useful for working on the quality of implementation of the treatment prior to the year in which outcomes would be measured.

Overall, our recommendation is that evaluators who are planning RCTs to measure the efficacy of technology-based interventions consider how analytics could be used to measure implementation fidelity against the program logic model and across all categories of fidelity. The low cost, timeliness, and objectivity of analytics make it a valuable new tool—which can supplement traditional interview, observational, and self-report measures—and can lead to better control of the expected contrast between conditions. This, in turn, can improve the quality of an efficacy trial.

References

Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Wiley and Sons, Inc.

Brown, M. (2011). *Learning analytics: The coming third wave.* EDUCAUSE Learning Initiative. Retrieved Nov. 2, 2013 from http://net.educause.edu/ir/library/pdf/ELIB1101.pdf.

Cordray, D. S. (2008). *Fidelity of intervention implementation.* Paper presented at the IES 2008 Research Conference.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*, 193-221.

Dron, J. and Anderson, T. (2009). On the design of collective applications, *Proceedings of the 2009 International Conference on Computational Science and Engineering*, Volume 04, pp.368-374.

Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Dusenbury, L., Branningan, R., Falco, M., & Hansen, W.B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237-256.

EDUCAUSE (2013a). *The NMC Horizon Report: 2013 Higher Education Edition*. Retrieved Nov. 5, 2013 from http://net.educause.edu/ir/library/pdf/HR2013.pdf.

EDUCAUSE (2013b). *The NMC Horizon Report: 2013 K-12 Edition*. Retrieved Nov. 5, 2013 from http://www.nmc.org/pdf/2013-horizon-report-k12.pdf

Fagan, A.A., Hanson, K., Hawkins, J.D., & Arthur, M.W. (2008). Bridging science to practice: Achieving prevention program implementation fidelity in the community youth development study. *American Journal of Community Psychology*, *41*, 235-249.

Feng, M. & Heffernan, N.T. (2007). Towards Live Informing and Automatic Analyzing of Student Learning: Reporting in the Assistment System. *Journal of Interactive Learning Research.* 18 (2), pp. 207-230. Chesapeake, VA: AACE.

Feng, M., Heffernan, N. T., & Koedinger, K. R. (2009). Addressing the assessment challenge in an online system that tutors as it assesses. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research, 19*(3), 243-266.

Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). Comparison of traditional homework with computer supported homework: Improving learning from homework using intelligent tutoring systems. *Journal of Research on Technology in Education (JRTE). 41*(3), 331–359.

Mowbray, C.T., Holter, M.C., Teague, G.B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation, 24*, 315-340.

Nelson, M C.; Cordray, D S.; Hulleman, C S.; Darrow, C L.; Sommer, E C. (2010). *A Procedure for Assessing Fidelity of Implementation in Experiments Testing Educational Interventions*. Society for Research on Educational Effectiveness. 2010.

Pardos, Z. A., Baker, R. S. J. d., San Pedro, M. O. C. Z., Gowda, S.M., Gowda, S. M. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, Leuven, Belgium. 117-124.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning*. IES practice guide (NCER 2007-2004). Washington, D.C.: National Center for Education Research. Retrieved February 28, 2008, from http://ies.ed.gov/ncee/wwc/pdf/practiceguides/20072004.pdf

San Pedro, M. O. Z., Baker, R.S.J.d., Gowda, S.M., & Heffernan, N.T. (2013). Towards an understanding of affect and knowledge from student interaction with an intelligent tutoring system. *Proceedings of the 16th International Conference on Artificial Intelligence and Education*, Memphis, TN.

San Pedro, M. O. Z., Baker, R.S.J.d., Bowers, A.J., Heffernan, N.T. (2013b). Predicting College Enrollment from Student Interaction with an Intelligent Tutoring System in Middle School. *Proceedings of the 6th International Conference on Educational Data Mining*, Memphis, TN. pp. 177-184

Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N., Razzaq, L., & Dailey, M. (2011). Improving K-12 homework with computers. In *Proceedings of the Artificial Intelligence in Education Conference,* Auckland, New Zealand.

U.S. Department of Education (2012). *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief*. Washington, D.C., 2012.

Acknowledgments

This material is based upon work supported by the Institute of Educational Sciences (IES) of U.S. Department of Education under Grant Number R305A120125. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES.

Where Are We Now? Research Trends in the Learning Sciences

Elizabeth Koh[1], Young Hoan Cho[2], Imelda Caleon[1], Yu Wei[1]
[1]National Institute of Education, Nanyang Technological University, 1 Nanyang Walk Singapore 637616,
[2]Seoul National University, South Korea, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742,
Email: elizabeth.koh@nie.edu.sg, yhcho95@snu.ac.kr, imelda.caleon@nie.edu.sg, yu.wei@nie.edu.sg

Abstract: Towards gaining a better understanding of the field of the Learning Sciences, this research investigates the research trends over 10 years. It also compares the Learning Sciences with the closely related academic fields of Educational Technology and Educational Psychology. A content analysis is performed on 5187 journal articles drawing from 12 top journals from 2003 to 2012. This content analysis was semi-automated and guided by an initial theoretical frame. The results reveal that research trends in the Learning Sciences have remained largely consistent except in the area of individual differences and affect, which has increased over the years. Key strengths of Learning Sciences include research on small group learning, inquiry, problem solving, argumentation, and mixed-methods. As the LS reflects on its state of practice, it should recognize that the field has achieved many research distinctives, yet, there are several opportunities for further research growth.

Introduction

The field of the Learning Sciences (LS) has evolved and blossomed internationally. It counts over 24 courses and programs around the world (NAPLeS). Over the past decade, new research topics and themes have emerged in the field including design studies, scaffolding, case-based reasoning, prior knowledge, and metacognition (Kolodner, 2004; Sawyer, 2008). LS researchers have been working closely with educators in schools, exploring new models of schooling, making explicit learning processes, as well as designing new methodologies and technologies to enable enhanced learning. As part of tracing its development and process, and enhancing its progress in the future, it is important to step back and reflect on the state of the field. Where are we now? What has been the focus of the LS? Are there certain research themes that have been dominating? What are the strengths and weaknesses of the field?

To gain some insight into the research trends of the LS, it is useful to compare it with related fields. The Learning Sciences draws from a wide spectrum of disciplines such as Cognitive Science, Educational Technology (ET), Educational Psychology (EP), Computer Science, and Applied Linguistics. The fields of ET and EP share a special bond with the LS as they all emphasize some aspect of learning. ET focuses more on the medium which learning occurs while EP, what occurs within the learner. It would be interesting to know whether they are differences between the LS and these two fields. Important distinctions could be drawn from the comparison, and theoretical and empirical contributions of the LS made explicit.

Moreover, in the field of the LS, there have been limited studies on its research trends. Several researchers have provided conceptual viewpoints of the state of the field. For instance, Sawyer (2008) synthesized four findings of LS research: 1) importance of deep conceptual understanding, 2) importance of learning connected and coherent knowledge, 3) learning authentic knowledge, and 4) collaboration. Besides conceptual viewpoints and insightful commentary, a useful method to examine the state of the field is through examining the corpus of articles published in the field.

Past research has examined published research trends through content analysis (Chang, Chang, & Tseng, 2010; Hew, Kale, & Kim, 2007; Hsu et al., 2012; Lee, Wu, & Tsai, 2009). The content analysis is usually conducted by categorizing journal articles in multiple dimensions. This is typically developed on the basis of literature reviews and/or a preliminary data analysis. For instance, Mitchell and McConnell III (2012) analyzed articles published in Contemporary Educational Psychology from 1995 to 2010 in terms of research topics, theoretical perspectives, research participants, and research methods. In addition, Hew, Kale, and Kim (2007) analyzed articles published in three journals related to ET from 2000 to 2004 in terms of research topics, research methods, data collection, and research settings.

Besides human content analysis, semi-automatic approaches to content analysis using text analytic software has been shown to be a viable method. These are more reliable than human coding and are relatively time-efficient (McKenny, Short, & Payne, 2012; Rotgans, 2012; Yu, Jannasch-Pennell, & DiGangi, 2011). For instance, Rotgans (2012) used semi-automatic content analysis to examine the trends of 10,168 articles in medical education research over 23 years. The research showed the viability of this approach and generated important themes useful for widening the scope of medical education research.

Through content analysis, we can identify what research topics, methods, research settings etc., are most frequently used in research articles and how their frequency has changed during a particular time period. For instance, Hew and colleagues (2007) found that articles about media study (41%) and psychology of learning (41%) are most frequently published in three ET journals from 2000 and 2004. Lee et al. (2009) found

that articles about student learning contexts were more published in three science education journals during 2003-2007 than during 1998-2002. Similarly, Rotgans (2012) found that there has been an increase in articles on communication skills training, clinical reasoning, faculty development, use of simulations, and student characteristics in six medical education journals from 1988 to 2010.

Towards gaining a better understanding of the field of the Learning Sciences, this research asks, what are the trends in LS over 10 years from 2003 to 2012? Secondly, how do these trends in LS differ from those in ET and EP? A research trend is the general direction that a field is heading towards. Trends will be identified by content analysis. A content analysis is performed on a dataset drawing from 12 top journals from 2003 to 2012. The quantity of articles that represent the research subtheme provides indication of the trends. This content analysis was semi-automated and guided by an initial theoretical frame of themes used in past research. The next section describes the methodology followed by the description of the results. A discussion of the findings is elaborated on and then integrated towards the end of the paper.

Method

Data Source

Journal article data from 2003 to 2012 was obtained from the Thomson Reuters Web of Science database. We selected four journals from each of the academic fields, LS, EP, ET. The twelve journals (refer to Table 1) are the journals with the highest impact factor in the three fields. We collected bibliometric information from original articles and excluded editorial materials, book and software reviews, and announcements. In total, 5187 journal articles were retrieved. Table 1 shows the breakdown of the papers from each journal. The International Journal of Computer-Supported Collaborative Learning was launched in 2006, although other LS journals regularly published articles from 2003 to 2007.

Table 1: Number of articles from the 12 Journals in the three fields

Field	Journal	No. of papers
Learning Sciences	(a) Journal of the Learning Sciences	140
	(b) International Journal of Computer-Supported Collaborative Learning	148
	(c) Cognition & Instruction	131
	(d) Instructional Science	290
Educational Psychology	(a) Educational Psychologist	186
	(b) Journal of Educational Psychology	657
	(c) Learning and Instruction	410
	(d) Contemporary Educational Psychology	270
Educational Technology	(a) Computers & Education	1502
	(b) Journal of Computer Assisted Learning	417
	(c) British Journal of Educational Technology	676
	(d) Educational Technology Research & Development	360

Analysis Method

All article titles, keywords and abstracts of the dataset were extracted and analyzed by semi-automated content analysis using the SPSS Text Analysis for Surveys 4.0 software. This software uses advanced linguistic algorithms to extract and classify key concepts from the text (IBM, 2011). The technology identifies the phrase, sentence and grammatical structures of content based on three main linguistic techniques: term derivation, term inclusion, and semantic networks. Based on pre-defined library rules, keywords and phrases are identified and grouped under a main concept. These concepts are then grouped into type patterns using semantic network techniques. For instance, the root concept “scaffolding” included related phrases and patterns such as “written scaffolds”, “peer scaffolding”, “computerized scaffolds + supported”, “scaffolding + effective” and “scaffolding approaches”.

One key advantage is that this technology allows researchers to extract and categorize key concepts reliably and consistently. It also reduces the time taken to categorize the content. However, human content analysis was still performed. Text analysis is an iterative process and extraction results were reviewed by the researchers according to the context of the text data. Categories were drawn from the data as well as derived from themes of past research. A preliminary data analysis of past content analysis frameworks revealed the following themes in Table 2. Based on this initial framework, relevant extracted concepts were identified. In addition, concepts that appeared frequently in the dataset were considered and fine-tuned as themes where possible.

Table 2: Initial Content Analysis Framework

Themes	Subthemes	Reference
Research topics	(a) Individual differences	Hew et al. (2007)
	(b) Learning processes	Hsu et al. (2012)
	(c) Instructional design and strategy	Lee et al. (2009)
	(d) Media study	Mitchell & McConnell III (2012)
	(e) Culture and community	
	(f) Research and evaluation methodology	
Research methods	(a) Quantitative	Hew et al. (2007)
	(b) Qualitative	Lee et al. (2009)
	(c) Mixed-methods	Mitchell & McConnell III (2012)
	(d) Review and meta-analysis	
	(e) Theoretical	
Research settings (participants)	(a) Laboratory	Hew et al. (2007)
	(b) Early childhood education	Hsu et al. (2012)
	(c) Primary school	Mitchell & McConnell III (2012)
	(d) Secondary school	
	(e) Higher education	
	(f) Informal context	
	(g) Virtual environment	

In total, from 5187 records, 1775 concepts were extracted. To better answer the two research questions of this study, we manually refined the software-generated categories by deleting irrelevant keywords, adding in new categories, editing the acceptable ones, and regrouping them. For example, the extracted data showed several articles focusing on "parent and family". We decided to identify this as our subtheme and included relevant extracted concepts such as "parent", "parental involvement", "father", and "mother". After the refinement process, 9 meaningful themes and 62 subthemes (subcategories) resulted. This formed the basis for further analysis.

From the SPSS Text Analysis for Surveys 4.0 software, a frequency count was generated. This denotes the number of articles that are classified into the subtheme. Even if an article has more than one extracted concept classified in that subtheme, it is only counted once. There are no extra counts of the same article in one subtheme. These larger themes were ordered by frequency from the highest to lowest. The results were then exported into SPSS Statistics 21 for further statistical analysis. For each subtheme, a series of three-way analysis of variance (ANOVA) was conducted to compare the paper counts for ET, EP and LS from 2003-2012. Whenever significant ANOVA results were obtained, Scheffe post-hoc test was conducted to compare the difference in paper counts between pairs of academic fields. To examine the trends over 10 years, we decided to use a ballpark period of 5-years to distinguish between earlier and later years, similar to Hsu et al. (2012). We conducted a series of t-tests to compare the paper counts for each subtheme of LS between period 1 (2003-2007) and period 2 (2008-2012). This process of refinement and analysis was similar to the steps by Rotgans (2012). Next, we examine the results of how themes changed over time (5-year interval) in the LS as well as how the themes were represented in the three academic fields.

Findings and Discussion

Table 3 below summarizes the results of the content analysis of the titles, keywords and abstracts of 12 journals from 2003-2012. This is ranked starting from the highest frequency count of the themes from the whole dataset.

Overall in the 3 fields, many of the articles fall into the theme of *research topics*, with 8232 articles examining a diversity of research topics from *media study* to *work and career*. Sociocultural practices such as *learning communities*, *culture* and *society* are also part of this theme of research topics. Within all the subthemes, there is a dominant subtheme on *media study*, with 3242 (62.5%) articles from the dataset. This is not surprising as various media forms from mobile devices to asynchronous discussion forums have been hot topics over the years in LS, ET and EP. This is similar to the findings by Hew et al. (2007). The subtheme of *student* in the category s*takeholder* is the second most examined area. The student being the main stakeholder in education research was found in 3108 studies (59.92%). At the tail end, there were fewer studies on *principal and leadership* among three fields (48 articles). This perhaps is due to the focus on the learner and the student in these education journals.

In the field of the LS, the *student* stakeholder is a primary focus. There were 478 articles or 67.42% of LS articles representing the student. Another major subtheme is that of *learning*, one of the educational activities. This was found in 348 (49.08%) articles. *Media study* was a close third, with 342 articles (48.24%) examining it. As for the tail end, *algorithms* were hardly examined, as just 1 article was identified. This is not surprising in LS as the focus is not on computer-related algorithms. *Principal & leadership* was also less of a focus, with 4 articles (0.56%) about it. This is an area of concern as school leaders can determine several

sociocultural practices in schools. In addition, *society* was a research topic that was hardly emphasized in LS with only 6 identified articles (0.85%). This is a surprising find as *society* provides a key context in the design and practice of learning. These infrequently covered subthemes highlight possible gaps in LS research.

Table 3: Result of the content analysis of the titles, keywords and abstracts of 12 journals from 2003-2012

Themes and subthemes	Comparison within LS[a]		Comparison among ET, EP, and LS			
	LS 2003-2007 (n=268)	LS 2008-2012 (n=441)	ET 2003-2012 (n=2955)	EP 2003-2012 (n=1523)	LS 2003-2012 (n=709)	Post-hoc tests[b]
1. Research Topics						
Media study	122(46%)	220(50%)	2655(90%)	245(16%)	342(48%)	ET>LS*, EP>ET*, ET≈LS
Academic achievement	19(7%)	38(9%)	242(8%)	472(31%)	57(8%)	EP>ET*, EP>LS*, ET≈LS
Learning environment	26(10%)	46(10%)	577(20%)	59(4%)	72(10%)	ET>LS*, ET>EP*, EP≈LS
Individual differences	7(3%)	26*(6%)	220(7%)	303(20%)	33(5%)	EP>ET*, EP>LS*, ET≈LS
Evaluation	13(5%)	28(6%)	470(16%)	40(3%)	41(6%)	ET>LS*, ET>EP*, EP≈LS
Learning strategies	11(4%)	24(5%)	359(12%)	94(6%)	35(5%)	ET>EP*, ET>LS*, EP≈LS
Memory	19(7%)	19(4%)	80(3%)	202(13%)	38(5%)	EP>LS>ET*
Learning outcomes	14(5%)	30(7%)	220(7%)	52(3%)	44(6%)	ET>EP*, LS>EP*, ET≈LS
Learning communities	19(7%)	32(7%)	217(7%)	22(1%)	51(7%)	LS>EP*, ET>EP*, ET≈LS
Curriculum	23(9%)	25(6%)	153(5%)	25(2%)	48(7%)	LS>EP*, ET>EP*, ET≈LS
Learning processes	9(3%)	23(5%)	133(5%)	30(2%)	32(5%)	LS>EP*, ET>EP*, ET≈LS
Culture	13(5%)	18(4%)	106(4%)	51(3%)	31(4%)	LS≈EP≈ET
Professional development	8(3%)	14(3%)	106(4%)	22(1%)	22(3%)	ET>EP*, EP≈LS, ET≈LS
Society	4(1%)	2(0%)	89(3%)	25(2%)	6(1%)	ET>LS*, ET>EP*, EP≈LS
Work and career	2(1%)	6(1%)	65(2%)	38(2%)	8(1%)	LS≈EP≈ET
2. Domains of Learning						
Knowledge	107(40%)	43(10%)	585(20%)	381(25%)	300(42%)	LS>ET*, EP>ET*, EP≈LS
Motivation	21(8%)	47(11%)	423(14%)	460(30%)	68(10%)	EP>ET*, EP>LS*, ET≈LS
Cognitive	61(23%)	95(22%)	333(11%)	368(24%)	156(22%)	LS>ET*, EP>ET*, EP≈LS
Skills	27(10%)	55(12%)	327(11%)	278(18%)	82(12%)	EP>ET*, EP>LS*, ET≈LS
Attitudes	10(4%)	19(4%)	315(11%)	69(5%)	29(4%)	ET>EP*, ET>LS*, EP≈LS
Beliefs	26(10%)	43(10%)	129(4%)	198(13%)	69(10%)	EP>LS>ET*
Affect	4(0.01%)	22*(5%)	178(6%)	148(10%)	26(4%)	EP>ET*, EP>LS*, ET≈LS
Metacognitive	12(4%)	21(5%)	62(2%)	82(5%)	33(5%)	LS>ET*, EP>ET*, EP≈LS
3. Stakeholder						
Student	176(66%)	302(68%)	1733(59%)	897(59%)	478(67%)	LS>EP*, LS>ET*, EP≈ET
Teacher	64(24%)	114(26%)	732(25%)	307(20%)	178(25%)	LS>EP*, ET>EP*, ET≈LS
Parent & family	3(1%)	6(1%)	66(2%)	112(7%)	9(1%)	EP>ET*, EP>LS*, ET≈LS
Pre-service teacher	6(2%)	15(3%)	127(4%)	23(2%)	21(3%)	ET>EP*, ET≈LS, EP≈LS
Policy-maker	5(2%)	5(1%)	109(4%)	16(1%)	10(1%)	ET>LS*, ET>EP*, EP≈LS
Principals & leadership	1(0%)	3(1%)	38(1%)	6(0%)	4(1%)	ET>EP*, ET≈LS, EP≈LS
4. Educational Activities						
Learning	134(50%)	214(49%)	1378(47%)	395(26%)	348(49%)	ET>EP*, LS>EP*, ET≈LS
Teaching	111(41%)	170(39%)	1196(40%)	413(27%)	281(40%)	ET>EP*, LS>EP*, ET≈LS
Assessing	19(7%)	19(4%)	261(9%)	68(4%)	38(5%)	ET>LS*, ET>EP*, EP≈LS
5. Research Settings (Participants)						
Higher education	45(17%)	75(17%)	826(28%)	255(17%)	120(17%)	ET>EP*, ET>LS*, EP≈LS
Secondary school	36(13%)	48(11%)	296(10%)	255(17%)	84(12%)	EP>ET*, EP>LS*, ET≈LS
Virtual environment	16(6%)	19(4%)	420(14%)	9(1%)	35(5%)	ET>LS>EP*
Primary school	11(4%)	25(6%)	250(8%)	134(9%)	36(5%)	EP>LS*, ET>LS*, EP≈ET
Early childhood	8(3%)	15(3%)	47(2%)	182(12%)	23(3%)	EP>ET*, EP>LS*, ET≈LS
Informal context	6(2%)	12(3%)	53(2%)	11(1%)	18(3%)	ET>EP*, LS>EP*, ET≈LS
6. Pedagogical Strategies						
Collaboration	69(26%)	134(30%)	556(19%)	61(4%)	203(29%)	LS>ET>EP*
Feedback	9(3%)	19(4%)	173(6%)	67(4%)	28(4%)	LS≈EP≈ET
Games	4(1%)	11(2%)	190(6%)	9(1%)	15(2%)	ET>LS*, ET>EP*, EP≈LS
Inquiry	33(12%)	54(12%)	85(3%)	36(2%)	87(12%)	LS>ET*, LS>EP*, EP≈ET
Simulation	9(3%)	15(3%)	155(5%)	12(1%)	24(3%)	LS>EP*, ET>EP*, ET≈LS
Small group learning	17(6%)	40(9%)	85(3%)	37(2%)	57(8%)	LS>ET*, LS>EP*, EP≈ET
Scaffolding	26(10%)	31(7%)	93(3%)	19(1%)	57(8%)	LS>ET>EP*
Reflection	10(4%)	16(4%)	102(3%)	20(1%)	26(4%)	LS>EP*, ET>EP*, ET≈LS

Themes and subthemes	Comparison within LS[a]		Comparison among ET, EP, and LS			
	LS 2003-2007 (n=268)	LS 2008-2012 (n=441)	ET 2003-2012 (n=2955)	EP 2003-2012 (n=1523)	LS 2003-2012 (n=709)	Post-hoc tests[b]
Argumentation	17(6%)	27(6%)	35(1%)	36(2%)	44(6%)	LS>EP>ET*
Modeling	8(3%)	10(2%)	66(2%)	28(2%)	18(3%)	LS≈EP≈ET
Problem solving	8(3%)	14(3%)	49(2%)	12(1%)	22(3%)	LS>ET*, LS>EP*, EP≈ET
Didactic teaching	3(1%)	6(1%)	51(2%)	16(1%)	9(1%)	LS≈EP≈ET
7. Research Method						
Quantitative	47(18%)	103(23%)	494(17%)	471(31%)	150(21%)	EP>LS>ET*
Qualitative	47(18%)	76(17%)	364(12%)	33(2%)	123(17%)	LS>ET>EP*
Theoretical, conceptual	23(9%)	34(8%)	195(7%)	57(4%)	57(8%)	LS>EP*, ET>EP*, ET≈LS
Mixed-methods	22(8%)	26(6%)	127(4%)	28(2%)	48(7%)	LS>ET>EP*
Review and meta-analysis	2(1%)	9(2%)	52(2%)	63(4%)	11(2%)	EP>ET*, EP>LS*, ET≈LS
Algorithms	0(0%)	1(0%)	65(2%)	7(0%)	1(0%)	ET>EP*, ET>LS*, EP≈LS
8. Epistemic Disciplines						
Science	65(24%)	135(31%)	311(11%)	182(12%)	200(28%)	LS>EP*, LS>ET*, EP≈ET
Mathematics	52(19%)	82(19%)	173(6%)	291(19%)	134(19%)	LS>ET*, EP>ET*, EP≈LS
Language	20(7%)	37(8%)	219(7%)	309(20%)	57(8%)	EP>ET*, EP>LS*, ET≈LS
9. Age of Learner						
Children	30(11%)	48(11%)	241(8%)	503(33%)	78(11%)	EP>ET*, EP>LS*, ET≈LS
Adult & lifelong learners	8(3%)	10(2%)	170(6%)	66(4%)	18(3%)	ET>LS*, EP≈ET, LS≈EP
Teenagers	3(1%)	9(2%)	38(1%)	149(10%)	12(2%)	EP>ET*, EP>LS*, ET≈LS

Note: [a]Only *p* values for the statistically significant *t*-test results were indicated for the comparison of LS paper counts for earlier years (2003-2007) and later years (2008-2012); [b]Results of Scheffe post-hoc test after significant results of a three-way analysis of variance was found on paper counts for the period 2003-2012; *p<.05.

What Are the Trends in LS over Ten Years (2003-2012)?

Overall, there were not many thematic differences between the earlier and later years of LS. The data revealed 2 significant differences in subthemes over the 2 stages. Development of the field of the LS over the 2 stages shows an emphasis on *individual differences* such as gender, learning styles, and age. As seen from Figure 1, there is a sharp increase of *individual differences* research compared to those of *evaluation* and *learning strategies* among the research topics. This suggests the recognition in the LS to examine sociocultural and larger contextual factors that affect learning.

Another finding is regarding *affect*, which examines feelings, moods, and emotions. This is increasingly focused on in the later years of LS. This suggests a broadening in the LS towards the understanding of the domains of learning. As illustrated in Figure 2, *cognitive* and *psychomotor (skill)* aspects are still frequently studied in the LS as compared to *affect*. However, the field has recognized the importance of the socioemotional aspects of learning too. We note that there seems to be a slight dip in the *cognitive* aspects of learning and this research area could be reaching a saturation point.

Nevertheless, the number of articles focusing on particular themes has been steadily increasing over the years. It suggests that, the LS as a field, is sticking to its roots. The next section helps us understand the LS research trends in relation to the closely related fields of EP and ET.

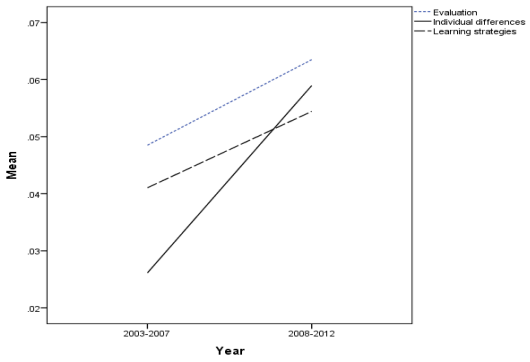


Figure 1. Trends in research topics

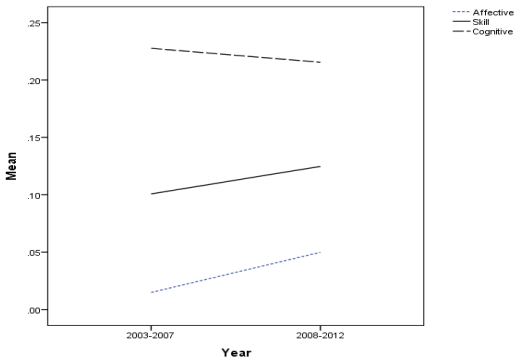


Figure 2. Domains of learning: affect, skill and cognitive

How Does LS Research Differ from EP and ET Research?

The results reveal many significant differences between LS, EP and ET. These differences highlight the strengths and weaknesses of each field. We examine the research topic first. Interestingly, the findings reveal that the LS does not dominate any particular research topic. The LS is similar to ET but stronger than EP in terms of *learning outcomes, learning communities, curriculum,* and *learning processes.* Focusing on learning end goals, being in learning communities of practice, designing the curriculum, and understanding processes of learning all suggest that the LS concerns itself very much in the practice of learning.

As for domains of learning, LS has focused more on *knowledge, cognitive*, and *metacognitive* aspects as compared to ET but similar to EP. In the LS, there is a clear emphasis on *knowledge* as seen from the many studies on knowledge building, knowledge creation and even knowledge transfer. As for *cognitive* and *metacognitive* aspects, these acknowledge LS's close ties with the Cognitive Sciences. It highlights the depth to which the LS explores learning, making visible thinking about thinking.

The LS holds *students* as their chief stakeholders, much more than the other 2 fields. Learning scientists seem to place importance on their students for it is them who are performing the learning. LS also emphasizes on *teachers* more than EP but similar to ET. This could be due to the tight interplay between teaching and learning in LS. What and how teachers teach affects how students learn. In ET teachers are important stakeholders too as they have certain control over the technologies that students use. LS and EP generally tend to emphasize naturalistic settings, where the classroom or course is the research setting and the teacher and student are the participants. However in EP, there is less emphasis on students and teachers as compared to the other fields possibly owing to their focus on laboratory testing with experiment subjects.

For educational activities, LS's strength is in *learning* and *teaching* but not *assessing*. There are a greater proportion of articles in LS on *learning* and *teaching* compared to EP as shown in Figure 3. This highlights the strong tradition in LS on these activities that directly contribute to enhancement of learning. *Assessing* is less emphasized in LS and suggests research opportunities in this area. Assessing is stronger in ET possibly due to the design and development of electronic forms of assessment.

In research setting, the LS has emphasized research on *informal contexts* and *virtual environment* as compared to EP. This is similar to ET. These highlights the many sociocultural understandings of learning, that it is not limited to the formal environment but to the informal and virtual contexts. Moreover, technological tools help to advance learning in these research settings. Interestingly, the other two fields have a larger proportion of articles on *primary school* research settings as compared to LS. EP also dominates the other two fields in *early childhood* settings suggesting a de-emphasis in the LS on these research participants.

In the 9th category, *age of learner*, a similar pattern emerges as EP dominates LS in terms of *children* and *teenagers*. It is possible that EP researchers with expertise in developmental psychology have carried out more research about individual differences in the development of children when compared to LS researchers. This suggests a gap in LS research for younger learners and LS researchers could delve into this area. As for pedagogical strategies, the LS shows its dominance in many different strategies*: collaboration, small group learning, inquiry, problem solving, argumentation,* and *scaffolding*. As can be expected, *collaboration* is a key

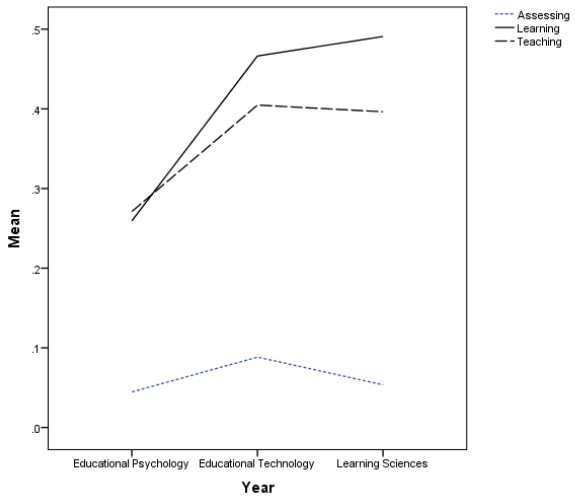


Figure 3. Educational activities among the three fields

focus in the LS. A journal on collaborative learning was specially launched. Similarly, the focus on *small group learning* is seen in LS compared to the other 2 fields. The *inquiry* approach is also a key strength of LS research.

The LS focuses much more on inquiry compared to ET or EP. This particular pedagogy systematizes how students discover and retrieve information. Entire curriculums have been developed around such pedagogies.

Problem solving, in particular real-world problem solving is another hallmark of the LS. Many curricular tasks are designed with problem solving as the basis to encourage students' reasoning, communication, and explanation skills. Many learning interventions in LS are centered on such problem-based learning approaches. *Argumentation* too has been clearly articulated in the LS with many LS scholars synthesizing the many steps and processes of such pedagogy. Another key strength of the LS is the subtheme *scaffolding*. The LS has a strong foundation of work on scaffolding compared to the other two fields. There have been many different theoretical developments of scaffolds from metacognitive ones to fading scaffolds and this pedagogical strategy has certainly contributed theoretically and practically to the field (McNeill, Lizotte, Krajcik, & Marx, 2006).

In terms of research method, *qualitative* and *mixed-methods* are more dominant in LS compared to EP and ET. This reinforces the field's desire for richness in insights and in-the-field design interventions. However, further opportunities exist for *reviews and meta-analysis*, which seems to be stronger in EP.

As for epistemic disciplines, *science* is the dominant discipline in LS. *Mathematics* is also emphasized more in LS compared to ET. However, *language* research is still a particular focus of EP as compared to LS. This could be an area that more LS researchers could go into.

Concluding Remarks

Where are we now? This paper has revealed data-driven research trends in LS over 10 years, as well as in comparison with ET and EP. The LS in these 10 years has remained rather consistent in research focus. Still, there is evidence of incremental changes in research themes, particularly in *individual differences* and *affect*. The increasing focus on *affect* is similar to the trend in science education research (Lee et al. 2009) which looked at the affective dimensions of science learning. It is also surprising that subthemes related to sociocultural practices have not increased significantly across the stages. Although many LS researchers agree that novel pedagogical strategies should be investigated in real contexts, it is hard to examine the roles of sociocultural factors in learning and teaching practices as well as learning outcomes. More LS research should explore dynamic interactions between individual learning activities and sociocultural characteristics of a community and develop new learning theories.

The comparison of LS with ET and EP reveals several distinctions of the LS. The LS has clear distinctives in the student as learner, pedagogical strategies: *collaboration, small group learning, inquiry, problem solving, argumentation,* and *scaffolding.* These lines of research should be continuously studied and capitalized on. Our data also reveals the dominance of LS in the epistemic discipline of *science*. As compared to EP, the LS has a greater emphasis on the research topics: *learning outcomes, learning communities, curriculum,* and *learning processes* and has focused on educational activities *learning* and *teaching* in *informal contexts* and *virtual environments*. As compared to ET, the LS is more dominant in the learning domains *knowledge, cognition* and *metacognition* and the epistemic discipline *mathematics*. Other strengths of the LS include its methodologies in *qualitative* and *mixed-methods*. It is important that LS as a new research field has unique topics, concepts, and theories, which are distinguishable from EP and ET.

We believe that there are areas for growth in LS. Research on assessment is an area of growth and there are some past LS research that has examined *assessment* for deeper understanding (Sawyer, 2008). Another topic that LS could focus on relates to *societal* aspects. Past LS has been very much classroom-based, but ecological and systemic perspectives could enlighten how LS work can be seen in a broader lens, and would have further implications at policy and societal levels. As mentioned, LS has been increasing in *affect* research over the 10 years. This is a good sign, as compared to the other 2 fields, LS still lacks behind them. Greater emphasis of the affective aspects of learning would be fruitful for a holistic understanding of the learner. Moreover, LS researchers might want to diversify from *science* and *mathematics* epistemic disciplines to explore *languages*. Lastly, the LS should also place greater emphasis on younger learners such as those in early childhood education to allow for a more developmental understanding of the learner.

This research is not without its flaws. The findings reflect only 10 years of the selected journal articles, and may not be an extensive gauge of research trends. Another limitation relates to the method where articles are classified based on extracted concepts through pre-defined rules. Articles could have been classified into only 1 category. Similarly, this method did not allow for articles to be coded for each research theme. To mitigate this, the authors examined articles which fell into only one category and identified further concepts or patterns to classify it into other subthemes. Lastly, the themes and subthemes identified in the study were dependent on the extraction process. If certain keywords were not present, it was difficult for the software to recognize it. This posed difficulties when the authors wanted to examine less common trends such as "scaling and translation" and this was removed from the final subthemes. Another problem arose for overly frequently words such as "design". For instance, many abstracts used phrases such as "in the design of the study". Due to the lack of sophistication in the rules of extraction, these articles were also categorized into the subtheme

"design studies". As there were too many non-relevant articles included in this subtheme, the authors had to remove it in view of having more accurate subthemes. We acknowledge that these categorizations can be further refined. Nevertheless, the existing themes and subthemes do reveal important findings of the state of LS research.

LS has its key strengths in research ranging from *small group learning, inquiry, problem solving, argumentation* to *qualitative* and *mixed-methods*. These strengths can be further capitalized on and deepened. This paper has also identified research trends that could be further examined such as *assessing* and *affect*. In addition, this study's methodology has showcased a reliable and relatively time efficient method of content analysis which can be further built upon. As the LS reflects on its state of practice, it should recognize that the field has achieved many research distinctives, yet, there are several opportunities for further research growth.

References

Chang, Y.-H., Chang, C.-Y., & Tseng, Y.-H. (2010). Trends of Science Education Research: An Automatic Content Analysis. *Journal of Science Education and Technology, 19*(4), 315-331.

Hew, K. F., Kale, U., & Kim, N. (2007). Past Research in Instructional Technology: Results of a Content Analysis of Empirical Studies Published in Three Prominent Instructional Technology Journals from the Year 2000 through 2004. [Article]. *Journal of Educational Computing Research, 36*(3), 269-300.

Hsu, Y.-C., Ho, H. N. J., Tsai, C.-C., Hwang, G.-J., Chu, H.-C., Wang, C.-Y., & Chen, N.-S. (2012). Research Trends in Technology-based Learning from 2000 to 2009: A content Analysis of Publications in Selected Journals. [Article]. *Journal of Educational Technology & Society, 15*(2), 354-370.

IBM. (2011). IBM SPSS Text Analytics for Surveys 4.0.1 User's Guide.

Kolodner, J. L. (2004). The Learning Sciences: Past, Present, and Future. *Educational Technology: The Magazine for Managers of Change in Education, 44*(3), 37-42.

Lee, M. H., Wu, Y. T., & Tsai, C. C. (2009). Research Trends in Science Education from 2003 to 2007: A content analysis of publications in selected journals. *International Journal of Science Education, 31*(15), 1999-2020. doi: 10.1080/09500690802314876

McKenny, A. F., Short, J. C., & Payne, G. T. (2012). Using Computer-Aided Text Analysis to Elevate Constructs: An Illustration Using Psychological Capital. *Organizational Research Methods*.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153-191.

Mitchell, A. W., & McConnell Iii, J. R. (2012). A historical review of Contemporary Educational Psychology from 1995 to 2010. *Contemporary Educational Psychology, 37*(2), 136-147.

NAPLeS. (2013). NAPLeS Programs, from http://isls-naples.psy.lmu.de/members/programs/index.html

Rotgans, J. I. (2012). The themes, institutions, and people of medical education research 1988–2010: content analysis of abstracts from six journals. *Advances in Health Sciences Education Theory and Practice, 17*(4), 515-527.

Sawyer, R. (2008). *Optimising Learning Implications of Learning Sciences Research*: OECD Publishing.

Yu, C. H., Jannasch-Pennell, A., & DiGangi, S. (2011). Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *Qualitative Report, 16*(3), 730-744.

Acknowledgments

This paper refers to data from the research project "Trends of Learning Research From 2003 to 2012: Analysis of Articles, Authors, and Institutions in Selected Educational Journals" (OER39/12CYN), funded by the Education Research Funding Programme, National Institute of Education (NIE), Nanyang Technological University, Singapore. The views expressed in this paper are the authors and do not necessarily represent the views of NIE.

Defining Success in an Alternative High School: Resources for the Reframing of Education

Gavin Tierney, University of Washington, 1100 NE 45th Street, #200, Box 354941, Seattle, WA 98105
gtierney@uw.edu

Abstract: This study looks at what it means to be successful in an alternative high school comprised primarily of adolescents who had been unsuccessful and/or marginalized in their previous schools. This paper uses ethnographic methods to focus on two students who were successful in the alternative school and the ideational resources those students used to participate in and reflect on the school. Two ideational resources are highlighted in this analysis – a critique of mainstream education and a focus on community education, providing insights into ways to re-engage students in school.

Major Issues Addressed

I sat across from Nancy, a 17-year-old student at Redwood High School, the alternative school where I had been doing research for the past three months. I concluded my data collection by asking Nancy my final interview question, the catchall: 'Is there anything that I haven't asked about that you think I should know?'

'Maybe," she replied, "just to reiterate that Redwood has a lot of problems, but it also saved my life and it has saved a lot of peoples' lives and it's still like one of the most beautiful things that I've ever been a part of.'

I include Nancy's response as an introduction to the problem space that I seek to explore in this paper. On a grand scale I seek to understand the process in which students see school as life saving and beautiful. The metaphoric saving of lives is something that is not easily measured. However, that does not mean that it should not be studied. If we are to stem the tide of high school dropouts in public education, it is important that we study the ways students are successful in all educational contexts, but specifically the contexts that help catch students who have previously been unsuccessful in school. This research is an attempt to explore what it means to be successful for students who have been previously unsuccessful and/or marginalized in school. Specifically, in this study, I seek to understand the experiences and perspectives of students who are now successful in an alternative school community. This includes both an analysis of how these successful students participate in school and how they reflect on what it means to be a new student to the school. By understanding the ways previously unsuccessful/marginalized students learn and become successful members of their school community, I hope to better understand ways to increase access for the students who remain on the margins of education.

In this paper I ask the following questions:

1) What does it mean to be successful in an alternative high school?

2) What resources do students who are successful members of an alternative high school use to participate and define success and membership in the school community?

Potential Significance

For this study, I pull from literature that looks at the resources that individuals use to construct their views of the world and their place in it (Barron, 2006; Nasir, 2012). Nasir identifies four components that help support students' identity development: ideational resources, material resources, relational resources, and the ability for students to put something of themselves into practices. In this paper I focus, in particular, on ideational resources, which Nasir describes as ideas about oneself, one's place in the world, and an understanding of what is valued, specifically which practices are valued in a community. I've chosen to use ideational resources because of how they guide students as they learn and participate in practice, guiding the types of people they are becoming. I hope to extend the work on ideational resources by looking at the ways outsider communities (such as alternative schools existing on the outskirts of mainstream education) use ideational resources to help the identity development, engagement, and success of youth.

There already exists a longstanding body of work that has studied how identities and forms of participation are produced and reproduced through social categorization in mainstream education and the informal contexts surrounding it (Crosnoe, 2011; Eckert, 1989; McDermott, Goldman, & Varenne, 2006; Willis, 1981; Wortham, 2005). While some studies such as Eckert, Fine (1991), and Willis have looked at resistant social groups, they are studied in contrast to adaptive social groups and the mainstream schools that help produce these groups. In her study, Eckert even mentions an alternative school populated by "burnouts", but maintains her focus on the mainstream high school. Each of this studies is invaluable, helping educators understand the processes of how social categories are produced in schools and the, potentially negative, impact

 © ISLS

they can have on students. By placing my study in an alternative school, I want to add to this literature by examining how students previously unsuccessful and/or marginalized can be successful. The alternative context is critical to this examination, in that alternative schools, while still conventional in certain ways, introduce different values, norms, practices, and resources. Pope (2001) introduced the idea of 'doing school', which involves the mindsets and practices of, superficially, what it means to be successful in mainstream schooling, including completing assignments, getting good grades, meeting assignment deadlines, and teacher-pleasing behavior. In part, by looking at what it means to be successful in an alternative school, I am exploring what it means to 'do alternative school'. Those students who are successful in mainstream school are those who understand the rules of the game of 'doing school'. Eckert would say these students are more often the jocks of the school, who have greater access to the resources of the school. If alternative schools are, in part, a safety net to help prevent students from dropping out of school, then it's important as educators to understand the rules of the game in alternative schools.

Theoretical Approaches

In this paper I have chosen to foreground the construct of ideational resources to understand a) the conceptual resources available to the students to construct their identities and b) the ways that success was conceived amongst the school community. As students who have not been successful in school enter into an alternative school their ideas about themselves, their ideas about the world, and their ideas about their place in the world may likely need to be renegotiated in order for them to be successful. Similarly, the ideational resources available in any educational setting says a lot about the context. For example, positioning students as college-going is an ideational resource for the student to think about their world, themselves, and their place in the world. Positioning students as college-going also provides insight into what is valued in the school, home, after-school program, etc. In many ways the ideational resources that are available to students reveal, at least in part, how a given community views the world and the available identities within that community. For this paper, then, I conceive ideational resources as socially shared and constructed concepts, categories, and values that help orient individuals to what it means to be a successful member of a community.

In order to think about the socially shared nature of ideational resources within a school I also use Wenger's (1998) idea of 'mutual engagement' to think about the ways in which the alternative school community used ideational resources to define success. Wenger describes mutual engagement as the negotiated meaning behind certain practices. In order to understand what it means to be successful in an alternative school I look at the ideational resources that community members engage with to participate and reflect on their participation in the school community. In considering mutual engagement, I also consider the constructs of 'joint enterprise" and 'shared repertoire' in mind as I analyzed observed practices and stories of the school. Joint enterprise is something on which there is a shared focus and value, negotiated in the moment. Wenger (1998) adds an historical element to the practices of a community with shared repertoire, describing them as resources created in a community over time through joint enterprise. These resources, he says, can include artifacts, stories, tools, and historical events. Together, mutual engagement, joint enterprise, and shared repertoire describe both what the community does together and their orientation to those tasks.

To understand success in an alternative school and what ideational resources students used to become successful in the school I also use Lave and Wenger's (1991) concept of central and peripheral membership and what Wenger (1998) calls old-timers and newcomers. While practices and values are constantly being negotiated and renegotiated amongst participants, newcomers typically enter into a community on the outside or periphery. Over time, as they become old-timers, they may move towards central membership in the community. They may also remain on the periphery or leave the community. For this study, in order to understand what it means to be successful, I focused on students who are successful old-timers at the school, while recognizing the need to additional research on newcomers and following students on their trajectories in alternative schools.

Methodological Approaches

This study is a case study of two high school students in an urban alternative high school community. Specifically, I followed the two students across classroom settings in one high school alternative program, focusing largely on the Advisory class that the two students shared. In the next sections I will discuss the research setting and participants, followed by sections focusing on the collection, analysis, and limitations of the data.

Setting and Participants

Redwood High School

The setting for this study was in an urban alternative public high school. In 1970, seeking an alternative to educational options offered by the school district, a group of students, parents, and teachers founded Redwood

High School. In the 2010-2011 school year, Redwood had 332 students and 25 teachers and support staff. In that same year the student body was comprised of 72% Caucasian, 7% Black, 2% American Indian, 8% Hispanic, and 7% Asian/Pacific-Islander. In 2010-2011 21% of the student body qualified for free or reduced lunch. In addition, in 2010 and 2011 Redwood scored higher on the state high school proficiency exam than both the district and state in Science, Reading, and Writing and in 2010 they were 2 percentage points behind the district average and one percentage point above the state average in Math. What is not captured in the school demographics and test scores is the high percentage of students who identify as GLBT. This information was reported in both student and teacher interviews and observations of the advocacy practices in the school. In addition, the school has been featured in local news stories on GLBT students. The GLBT population and the GLBT advocacy practices, such as gender-neutral bathrooms and course introductions where students say what pronoun they prefer, had implications on the ideational resources available to students at Redwood.

The setting for this study was chosen as an extreme case (Patton, 2003) when compared to the greater landscape of more conventional schools. Redwood High School caters primarily to students who, for one reason or another, have been disconnected or disenfranchised from mainstream education – this could be students who, while successful, self selected out of mainstream education or students who have had poor academic or social success in conventional schools. In addition, Redwood was chosen as a site because it was a formal high school community setting, but one that, based on school self-description, had a strong focus on integrating student choice, integrating students' out-of-school lives with their academic learning, and in which students' and teachers' roles were different than those in conventional schools.

My entry into Redwood High School was facilitated by previous informal meetings I had with principal and teachers at the school. Then, in 2011, I approached Trevor (an English Language Arts teacher who had been at the school for ten years) to see if his Advisory class would be a part of my study. At Redwood, Advisories are multi-aged (9^{th}-12^{th} grade) classes, made up of roughly 20 students. A core component of Redwood High School is its Advisory system in which every student has an Advisory that meets weekly and, in conjunction with Advisory, an Advisor with whom Advisees meet individually on a regular basis. Since Advisory was a core component of the program at Redwood and the focus of my research was on what it meant to be successful in the school, my study design began in an Advisory- first looking at the Advisory as the unit of analysis and, from there, identifying the case study students. Ultimately, Trevor left the decision of whether or not to participate in the study to his Advisory, inviting me in to present my case, a process that, while nerve-racking, I now wouldn't have had any other way. After three weeks of data collection in Trevor's Advisory class, I identified Nancy and Lori as focal students for my study.

Nancy and Lori

Nancy and Lori represented elements of both typical sampling and variation sampling (Merriam, 2009). They were typical in that both Nancy and Lori were seniors at Redwood High School, with both students considering themselves artists and members of their Advisory and of the Redwood community. In order to focus on old-timers at the school, I chose students who had been students at Redwood for multiple years and were experienced members in the Advisory group. As per my research questions, I sought to understand what the ideational resources were used through the perspective of students who were successful old-timers at the school. Both Nancy and Lori were Caucasian, came from middle class backgrounds, and spoke English as their first language. While Nancy, Lori, and the interviewed teachers all indicated that there were successful old-timers at the school who were male, female, and transgendered and who came from varied economic, cultural, and racial backgrounds, the majority of Trevor's Advisory was Caucasian. While there was a mix of males and females in the class, my primary focus for selection was to choose successful old-timers who had different participation patterns from one another.

Nancy and Lori had noticeably different ways of participating in class. During class and Advisory, Nancy was outgoing and, while not domineering, would often talk during class, sharing insights and jokes. In contrast, Lori was quiet and reserved and, while she participated in small discussion groups, she rarely spoke in front of the whole class. Nancy was a self-proclaimed language arts student, was socially active, and co-taught a course with one of the teachers for new students to Redwood. She was in her third year at Redwood, having come from seven years at a Montessori school before having a series of panic attacks in ninth grade and feeling she could no longer continue with school the way she had. Lori was in her fourth year at the school, coming to the school after attending a public middle school where she had some minor disciplinary experiences and felt as if the adults in the school did not care. She was nearly fluent in Spanish and considered herself an artist, musician, and poet. A final difference between the two was the leadership role that they took in the school. Nancy took on a visible leadership role, co-teaching a class for incoming Redwood students and helped collect student input on a school improvement initiative. Lori on the other hand talked about wanting to give back to the school and participated in school initiatives such as a spring break community service trip to South America, yet I did not observe her taking on leadership roles in the school.

After identifying Nancy and Lori as focal students, I began to train my lens on them during Advisory. In addition, in order to triangulate my data, I observed each of them in another more academically focused class. For Lori, I observed her in a film class taught by Trevor and, for Nancy, I observed a class that she co-taught with another English teacher, Doris, for students who were brand new to the school. Both of these classes were chosen with the help of Nancy and Lori and the goal of observing Lori and Nancy in contexts that differed from Advisory. In addition to classroom observations, I also attended one-on-one meetings that Nancy and Lori had with Trevor.

Data Collection

The analysis for this paper comes from three interviews with Nancy, three interviews with Lori, one interview with Trevor, one interview with Doris (Nancy's co-teacher), six observations of Trevor's Advisory class, three observations of Trevor's film class in which Lori was a student, three observations of Doris and Nancy's introduction to Redwood class (a sort of "Redwood 101" class), and Advisory-Advisee individual meetings between Trevor and Nancy and Trevor and Lori. In addition, I collected artifacts from the students, from the classrooms, and from the school. Throughout the data collection I kept an audio reflective journal, which I used as one form of triangulation during analysis.

I conducted three interviews with both Lori and Nancy. Each interview took between 45 and 100 minutes and was audio recorded. The interviews occurred either in school or at a local coffee shop. In the first interview I used a semi-structured and think-aloud protocols with a focus on probing the students' views of the school and their past and current experiences in school. In addition to the students, I interviewed Trevor (the students' Advisor) and Doris (the teacher with whom Nancy co-taught). I interviewed the teachers only after my observations and student interviews were complete, as I wished to capture the students' engagement with limited bias of the teachers. Teacher interviews were conducted primarily in order to triangulate data and included questions about the school and success at the school.

During my observations of the Advisory class, the Redwood High School 101 class, the film course, and the individual meetings, I shifted between acting as a non-participant observer and a participant observer (Merriam, 2009) depending primarily on invitation from the participants. I both took observational notes and audio recorded all observations. In addition to the audio recordings, I video recorded three of the observations of Advisory, focusing on the participation of the focal students. I chose to both audio and video record in order to create a record of how the focal students were participating during the class. My initial unit of analysis for Advisory was the joint enterprises or collective activities occurring during the class.

Approach to Analysis

There were four phases of data analysis in this study: Open coding, focused coding based on theory and research questions, writing analytic memos looking in and across codes and data, and generating hypotheses about student engagement. Prior to these phases, all interviews, student meetings, and observed classes were either video or audio recorded. Recorded interviews were then transcribed verbatim using Inqscribe software. Similarly, class and student meeting recordings were content logged, with pertinent sections being transcribed. Video, audio, and transcripts were analyzed using ATLAS.ti software.

During open coding I read and reviewed the entire corpus of data, selecting five specific interviews to open code (two of each focal student and the interview of the Advisor). During this phase, various codes emerged around topics such as 'listening', 'emotional response', 'pressure', 'defines education', 'love/like', 'hate', 'school work', 'anxious', 'personal history', and 'self-care'. All codes during this phase of analysis were generated from the participants' own words and actions.

During focused coding I developed and applied codes based on theory and my research questions. Specifically, these codes were based around themes such as ideational, material, relational resources, and self-expression in practice (Nasir, 2012), mutual engagement, joint enterprise, and shared repertoire (Wenger, 1998), and definitions of success.

In the third phase of data analysis I wrote analytic memos connecting the codes. In addition, during this phase I probed specific codes and the quotes and moments of participation associated with those codes. During this phase I also triangulated data across participants, interviews, and observations. An example of this was analyzing the ideational resources that the focal students engage in, triangulating interview data with the social practices they engage in during class.

In the fourth phase of analysis I looked across the codes and memos to generate hypotheses. Having generated these hypotheses, I looked across the data to test my assertions and identify any disconfirming evidence. Two particular ways that I sought out disconfirming evidence was looking closely at the observations to see if resources identified largely through interviews were present in practice. I also relied on teacher interviews to identify any disconfirming evidence, specifically to see if definitions of success and ideational resources were specific to the two case study participants. No strong disconfirming evidence was identified. As a final step I had fellow researchers look at the data to help confirm and modify my assertions.

Data Quality and Limitations

In this study I was interested in definitions of success through the perspective of old-timers at the school. Doing so allowed me to see what ideational resources successful students used, providing a model of success. In order to understand, in detail, students' perspectives and experiences I focused primarily on two specific students. Limitations of this study are that it did not include newcomers to the school and was a small sample. However, the data I have collected and triangulated across interviews and observations looks deeply at the ideational resources successful students used to participate and reflect on participation in the school. A final limitation of the study is that, as a former alternative school teacher, the experiences that the study participants spoke of were often similar to experiences that I myself faced as an alternative educator. I attempted to address this limitation through triangulation, peer review, and grounding myself in the words and actions of my participants.

Major Findings

During one of her interviews Nancy, one of the two focal students from this study, discussed the challenges of co-teaching a course for students who were new to Redwood High School.

> You, like, have to close up so much and build up so many shells in order to exist in other schools. At Redwood, the way you exist at Redwood in the most efficient fun way is to completely unzip, but that's really hard for people because we don't teach that at all. It's the opposite of what we teach, and so everybody figures it out in our own way and sometimes people never do.

Nancy presents the idea that students build up defenses to survive in mainstream education, but that in order to be successful at Redwood one needs to take down those defenses. In my analysis, I identified two categories of ideational resources that Nancy and Lori used to become successful at Redwood. One category of resources focused on the world outside of Redwood, specifically *critiquing mainstream education*. This ideational resource was presented as allowing the students at Redwood to view their past educational experiences as a product of the educational system and not their own failures as students. However, external critique extended to looking at the social injustices that exist in society and Redwood, as a part of society. The other category of ideational resources looked inward and was focused on the *community of Redwood*. The specific resources within this category included a new model of adult-student relationships, of school community, and of the content of school extending beyond only academic disciplines. These two categories of ideational resources were identified through the student participants discussion of their past experiences at Redwood, their views of newcomers to the school, and how they described and participated in Redwood during the study. In the remainder of the paper I will go into greater detail, explaining each of the major categories of ideational resources and how they relate to success at Redwood.

A Critique of Mainstream Education

For Lori and Nancy, existing in the community at Redwood was often framed in the context of other schools and students' previous education, specifically the ways in which conventional schools, including the schools they had previously attended, were constraining and uncaring environments. During my first interview with Lori, I asked her about middle school, before she came to Redwood. Lori was typically soft spoken, her voice calm and steady. Yet, when she started talking about middle school her voice got louder and she began to appear physically agitated as she told me about the school principal's indifference towards her. "I was just a kid," she said, "and they were an adult and they had nothing to do with me and they did not care". She they went on to recount episodes from middle school where the adults interacted with her as unsympathetic disciplinarians. At another point in the interview Lori analyzed her experiences in middle school. "I feel like there's a, there's just a certain like system of like certain things that like mainstream high schools want you to learn like their way and no other way." In these quotes mainstream schooling is positioned as uncaring and rigidly structured. Throughout Lori and Nancy's interviews it was clear that Redwood was viewed as existing outside of mainstream education. Lori described her anger during her time in middle school, anger that came bubbling up when she spoke about her middle school experiences, yet she was also able to abstract her experiences in middle school, lumping her middle school in with the rest of mainstream education.

In her interviews Nancy described the process of developing a critical framework of mainstream education, a process that can be difficult for newcomers to the school, especially when it involves one's emotionally turbulent educational history. "These kids fucking hate adults," she said, "They've hated adults for so long and just like have been let down by adults consistently for most of their lives." In this quote Nancy identified that newcomers often come to Redwood with a certain view of school and education, one in which adults are the problem. These students come to Redwood with potentially harmful ideational resources. It could be argued that the students also come with relational resources (friends), material resources (clothing, cigarettes), and ways to write themselves into school practices (ditching, non-participation), all which contribute

to their school-going identities. The ideational resources that newcomers have previously used and those present in the alternative school then may be in conflict.

Nancy discussed the transition from mainstream schooling to Redwood as "people wrestling with letting go of the structures that have been shoved into their bodies". Developing a critique of mainstream education was one way that Redwood tried to help newcomers in the process of 'letting go'. I saw parts of this process during observations the class Nancy co-taught for students new to Redwood. Through discussion, Nancy and her co-teacher, Doris, elicited students to tell stories of their own educational history in which they felt marginalized or constrained. Through these shared stories, the students engaged in a common problem of being constrained by past educational experiences. Additional resources were provided when Nancy had the students watch a short animated lecture by Sir Ken Robinson (RSA Animate, 2010). In this short video Sir Ken Robinson discusses the ways in which the current educational system follows a manufacturing model and restricts creativity. Nancy also asked the class to read a chapter by Alfie Kohn (2006) that critiqued the unexamined practice of homework. Kohn argued that homework was almost completely unnecessary and was actually harmful to students. After reading the text, the class discussed the rare times that they felt homework was valuable. Both of these instances sought to engage newcomers in a form of self-expression by metacognitively evaluating their education experiences and that of mainstream education and positioning Redwood as a different form of schooling. While both the video and the reading were material resources, they provided ideational resources for the types of things that were put in question at Redwood and, through that, the possibilities of being successful in school.

However, while ideational resources are offered to students as they enter into Redwood, the adoption of these resources is not the same for every student. Both Lori and Nancy identified the fact that students' personal histories influence the speed to which they take on the ideational resources such as a critique of mainstream education. "Redwood isn't enough of an all encompassing scenario to bring everybody out that, out of what their pushed into," Nancy said during one of her interviews, "I think that we do a good job with people who haven't been completely poisoned by the culture that we live in, but there're people who we can't reach." In another one of the interviews with Nancy I asked about the students in the class she was co-teaching. In her response she identified a sub-group from the class.

> Like, they've already got Redwood down and they're just going to keep getting awesomer and awesomer. Like they came to Redwood already Redwood kids and there's a certain amount of people who just do that. I was one of them. You just come to Redwood and you're already a Redwood kid and so you don't need to spend a lot of time becoming a Redwood kid. So they came and just opened up.

In this quote Nancy brought up the idea of alignment in that students come to Redwood with more or less experience engaging in the ideational resources and practices present at Redwood. Those students who have self-recognized experiences engaging in the ideational resources present at Redwood, those students Nancy described as 'already a Redwood kid', are immediately on an inbound trajectory towards becoming an old-timer in the school, quickly adopting the resources available at the school. Put another way, for some students the boundary between past educational environments and Redwood may be more malleable, blurring the line between newcomer and old-timer. However, students that are unfamiliar with the ideational resources present at the school may not as easily become successful at the school. This may result in students remaining on the periphery of the school or perhaps dropping out.

School as Community and Family

As I mentioned earlier in the paper, the ideational resource of critiquing mainstream education was directly connected to the fact that they were now attending a school that they saw as outside the mainstream. In both student interviews and classroom observations, Nancy and Lori referred to the school community as the 'Redwood Bubble', indicating that there were somewhat different norms inside versus outside the bubble. In a class discussion about school improvement that I observed, Lori advocated for more work in the community in order to learn about the norms outside of school and to promote what the school stood for.

While Lori and Nancy critiqued mainstream education and their own educational histories, they framed Redwood as an open and welcoming community that was rarely divided into cliques. Nancy described her thoughts when she was a newcomer to the school: "Redwood is the best place in the world, Redwood is so fantastic, it's social justice oriented, the community is completely accepting, like magnanimously beautiful." Similarly, Lori described her initial months at Redwood: "And I just feel like I knew every single person… like I was pretty much friends with everyone and that was pretty great." When asked about the elements of the school that contributed to their sense of community both students immediately referenced Advisory.

Advisory as an Entry Point

Logistically, Advisory at Redwood met once a week and was a time when students heard school wide announcements. However, it was also a time that the students and the Advisor simply spent time together – sharing humorous readings, group discussions about classes or movies, or eating bread that someone had baked. When asked what the goal of Advisory is, Lori said, "we just laugh, we just goof around and I'm always just laughing… we always get into the greatest conversations and I just feel like it's a great part of Redwood." At other points in the interview Lori described Advisory as "just kind of like a family within the community", helping students "get close to people and just like realize that school isn't just about work. You know, it's about community and interacting with other people that you wouldn't normally." In both of these quotes Lori described the ideational resource that Advisory provided, in which the roles of students and teachers, Advisory norms, and the forms of self-expression were based on laughing and existing as a small community. Advisory provided an example of how to interact with the world at Redwood, with old-timers modeling this process to newcomers. In a similar fashion, Nancy discussed the importance of Advisory for a new student to the school, describing it as "a safe space where you can be in a large group of your peers… and have it not be related to homework in any kind of context… Advisory is like where you ask questions when you're terrified."

Thinking back to earlier in the paper when Nancy brought up the idea of unzipping, it seems that the practices of Advisory acted as a resource to help students "unzip" and be successful at the school in non-academic terms. According to Nancy, Advisory provided a smaller community in which new students could become more successful members of the school, participating in the joint enterprises of laughing, talking, and asking questions about the opportunities and practices of the school. This then opened the potential for students to become successful members of the larger school community.

Teachers as Advisors, Advocates, and Friends

Advisors also acted as an important relational and ideational resource. Lori talked about the valuable role that Advisors play, saying: "It's just so important to have like that one person that you look up to there to help you through everything and not just school shit, like everything in your life." In this quote Lori talked about the value of having an adult who is involved in more than just schoolwork. Indeed, she had adopted the ideational resource in which having teachers involved in your life is prioritized. Nancy further identified the different roles of teacher and student when she said, "I'm really good friends with Trevor [her Advisor] I think, I like to think that, and I really care about him a lot." Similar to Advisory, Advisors at Redwood are very much a relational resource for the students. Advisors provided a positive relationship that helped students connect with the practices at Redwood. However, through the relational resources of Advisory and Advisors, students are also introduced to ideational resources that redefine what relationships look like in school amongst students and with adults. So, too, are students provided ideational resources that redefine how time in school is spent and the definition of school success. Within this ideational resource that Advisors provide, the role of a teacher and the role of a student shift and the boundaries of teaching are no longer confined to the school day or specific content. Nancy discussed the ways in which adults at Redwood presented themselves as relational and ideational resources. "The teachers really do a lot of work to like make sure that students know that they care about them and trust them. That's a really important thing at Redwood, that the teachers trust the students." It's important to note that, in Nancy's quote, teachers are not just relational resources by the nature of supporting the students, but that they trust the students as well. Again, this relational resource provided an ideational resource model of adult-student relationships.

Relevance to the Conference Theme

The themes identified in this paper provided insight into what it means to be a successful participant at Redwood. They also provide insight into the ideational resources available to students to become successful at the school and some of the processes of doing so. The ideational resources of critique and building community played different and complimentary roles in transitioning students to becoming old-timers to the school. The ideational resources promoted Redwood as being different from other schools, while also helping students simultaneously rewrite their past educational experiences and identities and creating new ones. Instead of viewing their prior educational experiences as a product solely of themselves or their mismatch to the setting, this process of renarration asked the students to adopt an awareness of the limitation of the structures of their previous educational experiences. Students were also pushed to consider school as a community and a family, a model of school relationship different than much of mainstream education.

It may be useful for learning environments to consider the resources of critique and community when trying to help student renegotiate their educational identities, especially with students who have built up defenses in mainstream schooling. Considering alternative schools specifically, this study indicates that alternative schools always exist in comparison and in contrast to mainstream schools and that the ideational resources provided in alternative schools are bound up in this alternative-mainstream relationship. It may be useful for alternative schools to leverage this relationship to mainstream schools, like Redwood did, developing

ways to help students re-engage in school. It may also be beneficial to consider the ways that new adult-student relationships, school community, and integration of students academic, social, and emotional lives played into the process of students becoming old-timers in an alternative school. Finally, in considering newcomers and old-timers, it may be useful to consider how these distinctions are in part a product of time in a community, but also students' familiarity with the resources, roles, relationships, and norms available in the alternative school.

This paper contributes to the conference theme of learning and becoming in practice by looking at some of the ways alternative school students learn to become successful students in school. In considering learning and becoming in practice, it is important to include the processes and practices of learning and becoming for students who have not been successful in school, in particular the ways that these students have re-engaged in school. Nancy and Lori identified that becoming a Redwood student entailed a process of renegotiating what in meant to learn and to be in school. By focusing on the ideational resources that these students used, I've presented potential ways to help more students become successful in school, increasing the possibility of, as Nancy put it, saving lives.

Endnotes

(1) All names of schools and participants are pseudonyms

(2) By 'unsuccessful' I cast a wide net, considering success in mainstream schools to be both academic and social in nature. By 'marginalization' I mean the social process of individuals feeling powerless in their own education. Often marginalization involves historically marginalized groups such as racial and ethnic minorities, youth labeled disabled, and youth identifying as lesbian, gay, bisexual, transgender, or questioning youth (LGBTQ).

(3) Willis does follow "the lads" into informal settings where they are the dominant culture and Eckert makes note of the working class settings that help inform the culture of burnouts.

References

Barron, B. (2006). Interest and Self-Sustained Learning as Catalysts of Development: A Learning Ecology Perspective. *Human Development, 49*(4), 193-224.

Crosnoe, R. (2011). Fitting in, standing out: Navigating the social challenges of high school to get an education. Cambridge: Cambridge University Press.

Eckert, P. (1989). *Jocks and Burnouts*. New York: Teachers College Press.

Fine, M. (1991). Framing dropouts: Notes on the politics of an urban public high school. Albany, N.Y: State University of New York Press.

Kohn, A. (2006). The homework myth: Why our kids get too much of a bad thing. Cambridge, MA: Da Capo Life Long.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge [England]; New York: Cambridge University Press.

McDermott, R., Goldman, S., & Varenne, H. (2006). The cultural work of learning disabilities. *Educational Researcher, 35*(6), 12-17.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.

Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences, 17*(2), 143-179.

Patton, M. Q. (2003). Qualitative research and evaluation methods. Thousand Oaks, CA: Sage Publications.

Pope, D. (2001). Doing School: how we are creating a generation of stressed out, materialistic, and miseducated students. New Haven, CT: Yale University Press.

RSA Animate. (2010, October 14). RSA Animate – Changing Education Paradigms. Perf. Ken Robinson [Video file]. Retrieved from http://www.youtube.com/watch?v=zDZFcDGpL4U.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, U.K.; New York, N.Y.: Cambridge University Press.

Willis, P. E. (1981). *Learning to labor: How working class kids get working class jobs*. New York: Columbia University Press.

Wortham, S. (2005). *Learning Identity: The Joint Emergence of Social Identification and Academic Learning.* Cambridge: Cambridge University Press.

"Case n' Point": Discovering Learning in the Nonce

Timothy Koschmann, Southern Illinois University, Dept. of Medical Education, tkoschmann@siumed.edu
Alan Zemel, University at Albany, Dept. of Communication, azemel@albany.edu
Michael Neumeister, Southern Illinois University, Dept. of Surgery, mneumeister@siumed.edu

Abstract: We present a single-case analysis of the taking of a stitch within a surgical procedure. All work carried out in the OR of a teaching hospital is held to exacting standards and scrutinized at all points regarding its adequacy. We examine how this kind of assessment is carried out with respect to a particular stitch taken by a surgeon-in-training in the course of a long and complex surgery. Our analysis offers an alternate way of talking about learning, treating it as an occasioned and interactional phenomenon. Learning, by this view, represents a special kind of performance done [1] publicly [2] for assessment and [3] with a displayed orientation to the next-time through. Implications for research in the learning sciences are developed.

Ways of Considering Learning

A father assists his daughter who is attempting to ride a bicycle on a playground. He encourages her saying, "It's okay, honey, you're *learning*." In another setting, an adult who is in the process of acquiring a second language engages in an exchange with a native speaker and says to himself, "I'm really *learning* this language!" Our interests are in ways that the term *learning* gets employed in these situations. What about the child's behavior in the first instance makes it recognizably learning? In the second, what is the adult doing when he says that he is learning? Is it just, in both cases, that they seem to be getting better at *X* or are they actually doing something when they say they are learning?

Psychologists talk about learning in terms of change over time. A classic definition was provided by Hilgard and Bower (1966) who stipulated:

> Learning is the process by which an activity originates or is changed through reacting to an encountered situation, provided that the characteristics of the change in activity cannot be explained on the basis of native response tendencies, maturation, or temporary states of the organism (e.g., fatigue, drugs, etc.) (p. 2)

Though much has been written since about how learning might be theorized (see Koschmann [2011] for a more detailed discussion of contemporary theories of learning), most current formulations conform to this definition at least to the extent they require the detection of change across situations. They may differ regarding how to characterize the changes, but some sort of change across situations appears to be criterial. Detection requires a "same-but-different analysis" (Koschmann, 2013) on the part of the observer, in that, in order to register as learning, the activity can't be so radically changed that it is no longer recognizable as what it is (i.e., maze running, doing maths, speaking French). Given this orientation to recognizable change, measurement becomes a preeminent concern. Educational psychologists, from Thorndike on forward, have designed their experiments such that different instructional regimens serve as treatment variables and performance measures serve as the dependent variable. But, operating under this paradigm, learning becomes doubly "occult" (Koschmann, 2002)"it cannot be seen on any particular occasion and it can never be observed directly. It is only known through its effects and it is only observable through the test instrument.

The question we would like to raise here is whether or not it might be feasible to treat learning as a concrete matter, as something available in the moment. Is it, in other words, investigatable as an occasioned matter? Is the little girl on the bike doing something that is recognizably and accountably learning? Can we, as in the example of the adult L2 learner, recognize learning when we are doing it? If we can answer these questions in the affirmative, it would suggest that psychology's formal definition of learning may not be the only way in which the phenomenon can be understood, that there are ways of recognizing learning as it is being produced. But this then opens into a larger question of just *how* this gets done and that is the question that motivates the current study. When complete, we would like to be able to say something about our common sense methods for recognizing and displaying learning. At the same time, we would like to be able to say something about how learning, so understood, is related to instruction. We see these inquiries as closely intertwined with the conference theme of "Learning and Becoming in Practice."

 © ISLS

Analysis

One might expect the operating room (OR) of a teaching hospital to be a "perspicuous" (Garfinkel, 2002, p. 181) setting for exploring how learning and instruction are done in situ. Here, a continuous supply of newcomers cycle through, leaving the program as certified practitioners. All work carried out in the OR is held to exacting standards and is carefully scrutinized regarding its adequacy. But, it is essential, for the purposes of training, to provide opportunities for trainees (i.e., surgical residents) to attempt technical elements of the procedure. These two, sometimes conflicting, agendas must somehow both be satisfied, but in a way that ensures that the patient receives the best possible care. Opportunities for practice can only be afforded if the elements fall within the present capabilities of the trainee and responsibility for making this determination rests with the supervising surgeon, the attending. Displaying and recognizing developing capacities, as a result, are organizational necessities within this environment.

Preliminaries

The study utilizes recordings from the SIU Surgical Education Video Corpus, a collection of recordings gathered over a dozen years at several surgical training sites. (1) We will present here a single-case analysis focusing on a fragment of interaction that occurred within a particular surgical procedure. The procedure was a Mastectomy and Free-Flap Breast Reconstruction. It required over nine hours to complete and involved three surgical teams. We will focus specifically on a single exchange during a critical part of the breast reconstruction between ATT, a highly-experienced plastic surgeon, and RES, an upper-level surgeon-in-training.

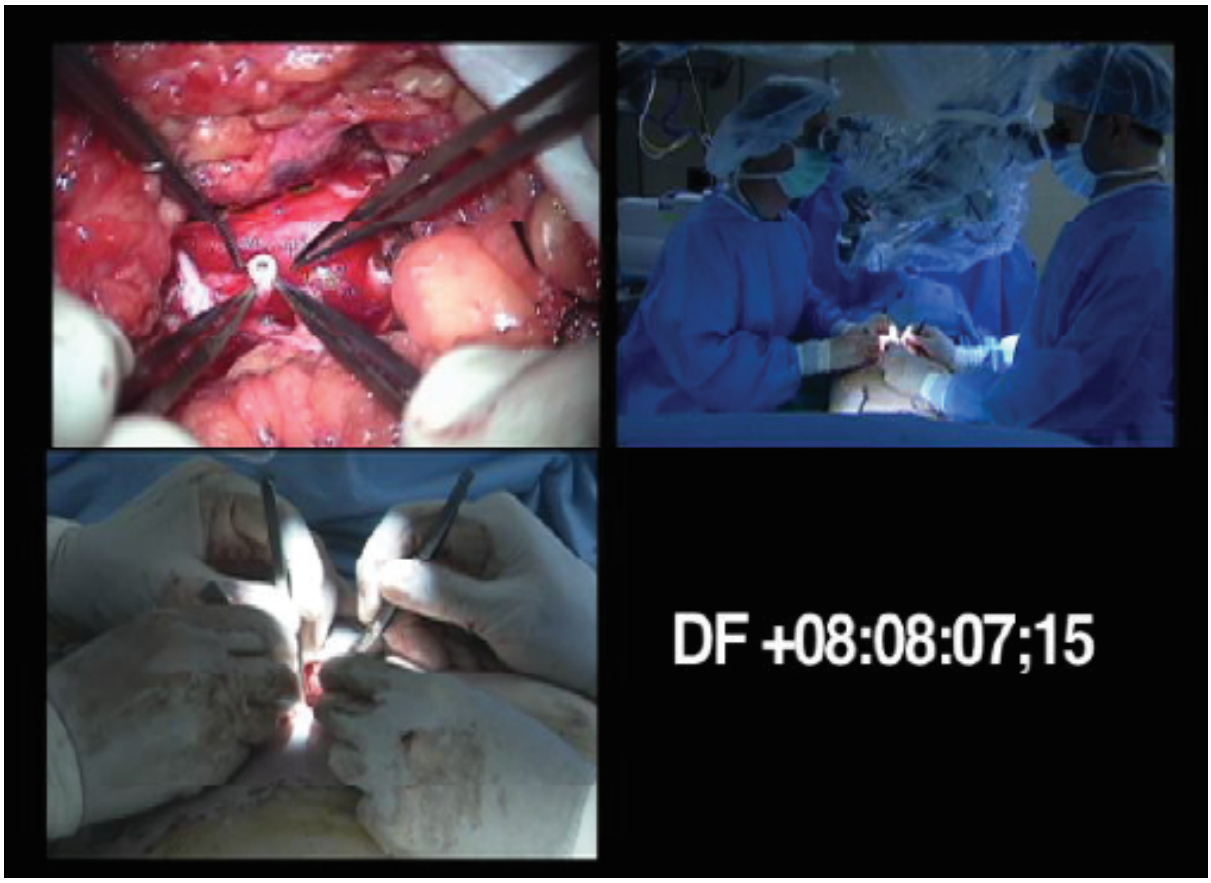


Figure 1. RES inserts the needle in the proximal vessel.

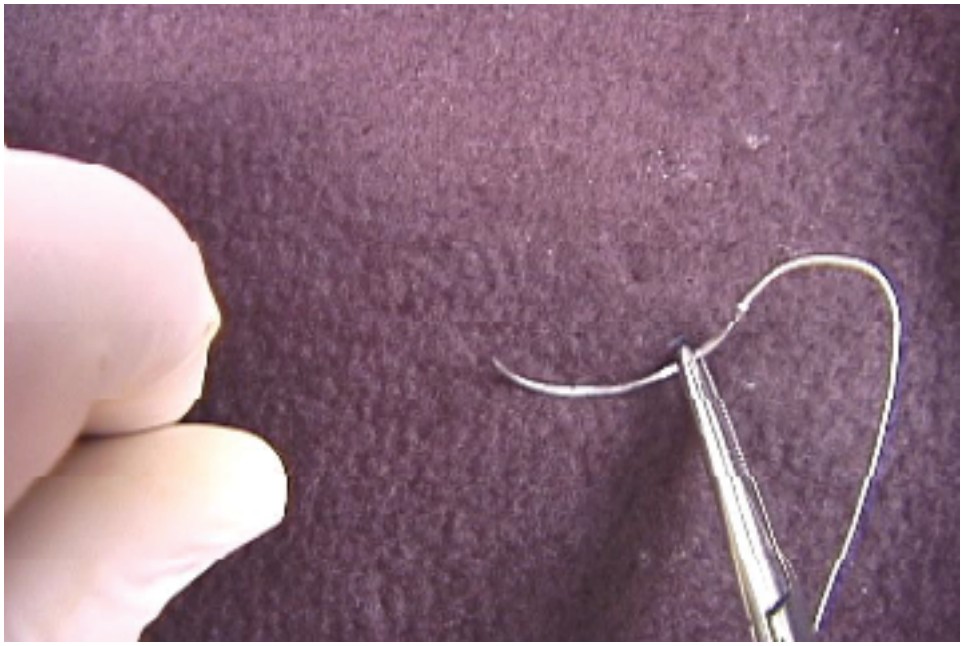


Figure 2. Curved needle grasped in a needle-holder.

In a free-flap procedure, the missing breast is reconstructed from tissue harvested from some other part of the patient's body, in this case the patient's belly. The transplanted tissue is referred to as the "flap" and its vascular bundle is referred to as the "flap pedicle." The most technically difficult part of a free-flap procedure is to re-connect the blood supply to and from the transplanted flap. The joining of two vessels is referred to as a 'vascular anastomosis'. Two anastomoses are required in a TRAM Free-Flap procedure, one to re-connect the primary artery and the other to re-connect the vein, known respectively as the arterio-arterial (A-A) and the veno-venous (V-V) anastomoses. An anastomosis of two vessels is formed by making a series of stitches, enough to ensure that there are no visible gaps and that a tight seal is established. In the surgery observed, a dozen sutures were required to complete the A-A anastomosis. Our analysis will focus on the production of the first.

"Case 'n Point"

As we join the scene, ATT and RES stand on opposite sides of the operating table. A two-person, stereo-microscope is suspended over the table between them. Some vascular surgery can be done with minimal magnification, but the vessels involved in a free-flap anastomosis are relatively small and microsurgical technique must be employed. Because of the diminutive size of the structures with which they are working, everything under the microscope is manipulated using specialized instruments. A common tool is a kind of forceps resembling a long-handled tweezers referred to colloquially as "pick-ups." The surgeons must steady their hands by bracing them and resting them on the patient's body. As shown in Figure 1, both surgeons work with a tool in each hand. Appendix A represents the talk and some of the visible action that took place during the production of the first and second sutures of the A-A anastomosis. It was prepared using the standard transcription conventions of Conversation Analysis (CA). (2) The transcribed segment occurs shortly after the completion of the V-V anastomosis which was constructed by ATT with RES assisting.

Producing serviceable sutures is one of the first skills acquired by a surgeon. There are many kinds, the simplest being the interrupted suture. The procedure for making one involves hooking a curved needle through the respective edges of the two sections of tissue to be joined or "approximated." For an instrument tie, the needle is held using a special-purpose needle holder (see Figure 1). "Taking a bite" with a needle holder is a two-step process (Anderson & Romfh, 1980). The needle-holder is commonly clamped to the middle of the needle, so when the needle is inserted, it can only be pushed to this point. It must then be released and pulled through from the opposite side. This is repeated for the second piece of tissue, the one to be approximated to the first. Having pulled the needle through both sections, the trailing segment of suture becomes the working end for knot-tying purposes, while the thread attached to the needle becomes the standing end. They are joined using a "surgeon's knot", basically a modified square knot. It consists of a loop of fixed perimeter and two or more "throws" or wrappings of the loose ends (Edlich, 2008). The suture is completed by snipping off the free ends or "ears" of the joining strand.

As the fragment begins, we find RES issuing a directive to the scrub nurse to provide a needle-holder loaded with a needle and 9-0 suture. Unless otherwise specified, when an instrument is requested, the person

issuing the request is the expected recipient and that is the case here. By calling for the needle-holder, RES positions himself as the party responsible for performing the next item of business, the joining of the patient's truncated internal mammary artery to the dissected artery within the flap pedicle. The two surgeons individually prepare for the joint task ahead of them—ATT readying the vessel ends and RES organizing the needle with suture attached. The vessels must be lifted and supported while the stitches are being taken. Also, to successfully and precisely place the needle, the artery wall must be supported from the other side. As assistant, all of this work falls to ATT. As they begin, he inserts and expands a pick-up inside the distal (i.e., on the flap-side of the finished anastomosis) vessel (lines 6-7), giving RES better control of the needle placement. RES pushes the needle halfway through the arterial wall (lines 8-9), releases the needle-holder, and then re-uses it to draw the needle the rest of the way through (lines 10-11).

Next, a decision must be made about where to have the suture pass through the other, the proximal, vessel. This planning work takes the form of an insertion sequence embedded within the unfolding project to produce the first stitch. Misalignment of the two segments could lead to twisting of the anastomosis that might produce problems later. RES consults ATT before proceeding (l. 13). His query comes with a candidate answer embedded, displaying his ability to independently make the necessary judgment. It is accompanied by a swiping gesture performed with the needle (lines 14-15). The gesture, passing over the lower lip of the artery, is precisely coordinated with his enunciation of "that's" (cf., Hindmarsh & Heath, 2000). ATT first stretches the distal segment to line it up with the proximal. The movement serves both as a visual test of alignment and as a mediated point (lines 18-19). RES then proceeds to insert the needle at the identified position.

Having come to concordance on where the needle insertion should be made, RES proceeds, but there is another question about placement and this one pertains to how far back from the tissue edge the needle should be set. This judgment RES makes without consultation (l. 21). Immediately after RES sets the needle in the second (proximal) vessel, ATT issues a double directive pertaining to [1] the next stitch ("Take the next one bigger bites") and [2] the needle set or "bite" just accomplished ("take tha' one (0.6) j's a little bit bigger"). (3) The latter is produced without an accompanying gesture.

The attending's turn begins with "Take the next one." Though what follows pertains to the current bite ("tha' one"), RES may only be attending to the first part. Before ATT's turn is complete, RES proceeds to pull the needle through the proximal segment (l. 23) before issuing a receipt token (l. 27) to ATT's eventually completed turn.

His receipt is ambiguous—is he responding to the first directive, the second or both? Rather than immediately withdrawing the needle and repositioning it, he draws up on the suture, tugging the two artery segments together (lines 29-30). This constitutes a practical test of the suture in progress — both with regard to its strength and the alignment of the two segments. Unfortunately, when tension is placed on the suture the second bite fails (l. 31). ATT registers the problem using a non-lexical expression of dismay (l. 32). RES's reply (l. 34) relates the problem to the issue previously raised by ATT.

Instructing and Learning in Interaction

What is it about this brief fragment that offers an impression of instruction going on? What about it suggests the possibility of learning? In a now classic paper on the nature of instructional talk, Mehan (1979) offered two hypothetical exchanges, one that went like this:

> Speaker A: What time is it, Denise?
> Speaker B: 2:30
> Speaker A: Thank you, Denise.

and another that went like this:

> Speaker A: What time is it, Denise?
> Speaker B: 2:30
> Speaker A: Very good, Denise.

The two sequences differ only in their third turns. In the first, we have a receipt and acknowledgement of the information offered, in the second, an evaluation of the information provided, suggesting in Mehan's terms that the question had a "known answer." The third turn in this way produces what came before it as an assessable (and potentially correctable) performance. But it does more—it serves to establish what will count as accountably correct performance. (4) This, then, becomes, what we have referred to elsewhere (Zemel & Koschmann, in press) as, a "learnable." At the same time, it assigns different epistemic roles to the parties, establishing one as knowledgeable with respect to the matter in question (the *instructor*), the other as standing in need of instruction (the *instructee*). These things are the earmarks of an instructional organization.

We can see a semblance of this in the analyzed fragment. The functional equivalent of the instructor's third turn can be found in ATT's double directive begun in l. 22 and completed in l. 25. Coming directly after the insertion of the needle in the proximal vessel, the attending's second directive calls for a correction to be implemented by RES. Other-correction is usually dispreferred in non-instructional talk (Schegloff, Jefferson, & Sacks, 1977) and the presence of correction here is part of what gives this an instructional character. This correction-initiation in this particular position accomplishes exactly the things we described with respect to Mehan's "What time is it, Denise?" example— it positions the parties in certain epistemic roles and produces, for current purposes, a normative order for bite setting. (5) Strictly speaking, it has some differences from the Mehan example, as well. Mehan's example involves positive assessment in the third turn, where here we have a negative assessment. In the classroom, we might find direct correction in the third turn rather than correction-initiation (but see McHoul, 1990). Nonetheless, we have no trouble recognizing this sequence as instructional. Indeed, it is more than instructional, it can be heard as a form of caution, though what it is cautioning against is left to the recipient to work out.

But if we see instruction here, do we also see learning? By 'learning', of course, we are referring not to the operationalized construct employed by educational psychologists, nor are we talking about some kind of hypothesized mental event. We are orienting instead to something that can be seen and heard within the participants' unfolding interaction. The instructional sequence consists of an elaborated performance begun at l. 8 and continued to l. 21 carried out with considerable assistance from ATT. After ATT's appraisal at lines 22 and 25, RES supplies a nominal uptake token (l. 27), but this, at best, represents an avowal of understanding, not a demonstration. Indeed, he does not make an immediate move to correct the faulty needle placement in the proximal segment. The subsequent failure of the stitch, in effect, serves as its own practical assessment. We see the sequence produced as learning with RES's "Case n' point" (l. 34). In accounting for their unfolding work in just this way, he links the failure of the stitch to ATT's prior caution. He formulates their current situation as a case of a stitch production in which one of the bites allowed too little cuff and, as a consequence, failed. In a way that his earlier receipt token (l. 27) did not, this formulation concretely demonstrates his understanding of the ATT's warning and "witnessably" (Rawls, 2002, p. 51FN) produces this as a learning sequence. Like the instances of instruction produced earlier, learning also entails assessment, but in this case it is self-appraisal by the learner that is critical, rather than other-assessment (and correction initiation) by the instructor. It is RES's treatment of the local occurrence as documentary evidence of a general principle, i.e. as a "case," that accomplishes an orientation to 'a next time through'. In this way, our analysis reveals a different way of thinking about learning as an occasioned and interactional phenomenon. It represents a special form of performance done [1] publically [2] for assessment and [3] with a displayed orientation to the next-time through.

Discovering Learning in the Nonce

Just as Wittgenstein (1958a, 1958b) used posed examples as vehicles for exploring particular philosophical questions, we can utilize the analyzed episode as a "propaedeutic case" (Garfinkel, 2002, p. 75) for thinking through what we take learning to be. Our analysis makes two important contributions in this regard. First, the Thorndikean tradition in educational psychology rests on the presumption that learning cannot be seen, that it is an "occult" phenomenon (Koschmann, 2002). With this analysis we demonstrate that it can in fact be observed in the course of its production, at least within this attested example. This is a significant finding.

Second, we saw in this analyzed example something interesting about the relationship between instruction and learning. While it is often assumed that one brings about the other, we see in this episode that the practices whereby learning is produced can, in at least some cases, be displaced from the practices whereby instruction is produced. This is not to suggest that the two forms of action are independent. Indeed, in the case analyzed, RES's 'learning' displayed his appreciation of the matter that he and ATT had jointly produced as 'learnable' within the instructional sequence. So the learning was sequentially tied to the instruction that came before, though the practices through which the instruction was produced were analyzably distinct from those that constituted the learning. That the two can be examined separately is a special feature of instances of instruction and learning like the one presented here.

Though our findings call into question a basic presupposition of the Thorndikean program, we would not like to suggest that the program itself is thereby rendered invalid or dismiss its accomplishments out-of-hand. Thorndike sought to put the study of learning on a scientific footing and, to do so, he imposed certain requirements on how the phenomenon could be approached. He believed that learning was inherently tied to assessment. He was, in this way, appealing to a common sense understanding of what it means to learn something, that we often impose checks or tests on ourselves and others to see whether learning has occurred. In our analysis we too found that appraisal and assessment are integral to both learning and instruction when they are studied in unfolding interaction. So assessment in some sense appears to be key to how we understand learning in all cases, but we may differ in terms of where we proceed from there. One part of the belief structure that produces learning as occult, is the belief that learning is only appreciable as change over time. Again, this appeals to our everyday sense of what it means to learn, but at the cost of making it impossible to locate the

phenomenon within actual unfolding conduct. In the surgical procedure from which the analyzed excerpt came, it would have been possible to examine RES's subsequent stitches for evidence that he had 'learned' from the "Bigger bites" lesson. Such an analysis would not necessarily differ from the ways in which we sometimes apply the term 'learning' in everyday circumstances. Both approaches require a "same-but-different analysis" (Koschmann, 2013). But we clearly need more empirical research into how these different conceptualizations, learning-as-change-over-time and learning-as-occasioned, are related.

Historically, educational research has advanced by axiomatizing learning, by stipulating from the outset what learning might be. We are not taking issue specifically with any of these formulations, but are instead calling into question this general way of doing business. If we are to achieve an empirical science of learning, ought that not engender some foundational inquiry into its central phenomenon, into the nature of learning itself? It is essential within any scientific enterprise to strike agreements pertaining to what it is that we are undertaking to study. But when it comes to learning, we seem to be locked into endless cycles of definitional propagation. Worse, we have, at least in some cases, adopted definitions that render the phenomenon of interest off limits to direct study. The current paper represents an effort to depart from this tradition. It could be characterized as an inquiry into what people do when they describe themselves as learning. Rather than stipulating *a priori* what it might be for the purposes of our investigation, we have sought to locate it within our empirical materials. We have taken it up as an endogenous matter and have sought to return it to the site of its production. In so doing, a whole new world has opened before us, one ripe for further exploration.

Endnotes

(1) Signed consent to record a surgical procedure is sought from the patient upon admission to the hospital. Advance consent is secured from all other participants (i.e., attendings, fellows, residents, medical students, staff) prior to recording in the OR. The consent forms for the participants and patients are associated with a collection protocol approved by the institutional review board (IRB). For purposes of confidentiality, recording is only begun after the patient has been completely draped. Also, all proper references (e.g., patient names, names of practitioners and institutions) appearing on the recording are redacted prior to study. To actually do research with the materials in the corpus, investigators must submit a second protocol, a use protocol, with the local IRB.

(2) The full set of conventions is described in Jefferson (2004). In brief, special brackets are used to mark the onset of overlap between transcribed elements (i.e., turns at talk or other transcribed conduct). Numbers enclosed in single parentheses represent periods of silence measured to a tenth of a second. Standard punctuation marks such as periods and question marks are used to denote delivery with falling (or rising) intonation. Colons are used to display sound stretching. Text enclosed between degree signs represents talk delivered at diminished volume. Annotations supplied by the transcriber are enclosed in double parentheses. These are most often used to describe visible conduct occurring in conjunction with the talk. The column appearing on the left side of the transcript presents the times, measured in hours, minutes, seconds, and frames, at which the actions, either talk or embodied conduct, were initiated. Line numbers are added on the far left to simplify reference in the text.

(3) Anderson and Romfh (1980) write:

> A minor well-conceived expansion of vocabulary will allow a surgeon to communicate with his helpers with less misunderstanding. It is an ambiguous instruction, for example, to tell an assistant placing a stitch to "take a bigger bite." Such a request could mean: make more progress between stitches; make a wider cuff; or take a thicker cuff of tissue. Planning a vocabulary that will allow you to communicate by non-ambiguous, specific instructions will greatly facilitate coordination between surgeons and assistants. (pp. 178-179)

It would appear that surgeons must struggle, just as we as analysts do, to find the right words to describe these practical matters.

(4) "What these direct instructional sequences yield, and what they are posed to yield, is something like accountably correct answers, and, by implication, knowledge and competence" (Macbeth, 2004, p. 704).

(5) It is important to note that it is not simple assessment that marks an exchange as recognizably instructional, but rather assessment coming on the heels of an assessable performance, the assessment and the assessable performance being reflexively related. It should also be noted that we are not claiming that this is *only* way in which instruction can be enacted. There may be and likely are any number of other organizations yet to be uncovered and documented.

References

Anderson, R. M., & Romfh, R. F. (1980). *Technique in the use of surgical tools*. New York: Appleton-Century-Crofts.

Edlich, R. F. (2008). *Surgical knot tying manual* (3rd ed.). Norwalk, CT: Covidien AG.

Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Lanham, MD: Rowman & Littlefield.

Hilgard, E. R., & Bower, G. (1966). *Theories of learning* (3rd ed.). New York: Appleton-Century-Crofts.

Hindmarsh, J., & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics, 32*, 1855-1878.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13-31). Amsterdam, Netherlands: John Benjamins Publishing.

Koschmann, T. (2002). Dewey's critique of Thorndike's behaviorism. Paper presented at the American Educational Research Association, New Orleans, LA.

Koschmann, T. (2011). Theorizing practice. In T. Koschmann (Ed.), *Theories of learning and studies of instructional practice* (pp. 3-17). New York: Springer.

Koschmann, T. (2013). Conversation analysis and learning in interaction. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1038–1043). Oxford, U.K.: Wiley-Blackwell.

Macbeth, D. (2004). The relevance of repair for classroom correction. *Language and Society, 33*, 703-736.

McHoul, A. (1990). The organization of repair in classroom talk. *Language and Society, 19*, 349-377.

Mehan, H. (1979). "What time is it, Denise?" Asking known information questions in classroom discourse. *Theory into Practice, 18*, 285-294.

Rawls, A. W. (2002). Editor's introduction to Harold Garfinkel's *Ethnomethology's program: Working out Durkheim's aphorism* (pp. 1-64). Lanham, MD: Rowman & Littlefield.

Schegloff, E., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*, 361-382.

Wittgenstein, L. (1958a). *The Blue and Brown Books*. New York: Harper & Row.

Wittgenstein, L. (1958b). *Philosophical investigations* (G. E. M. Anscombe, Trans. 2nd ed.). Malden, MA: Blackwell.

Zemel, A., & Koschmann, T. (forth.). “Put your fingers in there”: Learnability and instructed experience. *Discourse Studies*.

Appendix A

Excerpt 1: First stitch (#04-010)

```
 1  08:06:27;07 RES:  I'll take the nine oh
 2  08:06:28;18       (45.1)
 3  08:06:53;04       ((ATT organizes the distal artery segment using pick-
 4                    ups in his r. and l. hand))
 5  08:07:13;21 ATT: °°u:::h°°
 6  08:07:31;24       ((ATT inserts pick-up into distal segment opening it
 7                    in preparation for receiving the needle))
 8  08:07:32;25       ((RES inserts needle through distal segment using
 9                    needle-holder in r. hand))
10  08:07:41;10       ((RES pulls needle through distal segment with
11                    needle-holder))
12  08:07:48;26       ((RES loads needle in needle-holder))
13  08:07:54;07 RES:  I'think ⌈that's about the bottom don't you?
14  08:07:54:12               ⌊((RES swipes needle along lower lip of
15                    proximal segment))
16  08:07:55;26 ATT:  I think that's pretty close to being yea::h right
17                    (0.6) ⌈right (0.2) ther::e
18  08:07:57;10             ⌊ ((ATT gestures with stretched end of the
19                    distal segment))
20  08:07:59;09       (13.0)
21  08:08:07;15       ((RES inserts needle into the proximal vessel))
22  08:08:11;28 ATT:  Take the next one ⌈bigger bites and take tha'one =
23  08:08:12;21                         ⌊ ((RES pulls needle through with
24                    pick-up))
25  08:08:14:00 ATT:  = (0.6) j's a little bit bigger
26  08:08:15;11       (0.4)
27  08:08:15;23 RES:  Kay.
28  08:08:16;00       (1.6)
29  08:08:16;14       ((RES tightens suture, drawing the two artery ends
30                    together))
31  08:08:16;27       ((suture tears out of the proximal vessel))
32  08:08:17;17 ATT:  °nyeh°
33  08:08:17;24       (0.8)
34  08:08:18;19 RES:  Case 'n point
35  08:08:19;15       (7.3)
36  08:08:20;26       ((RES re-loads needle into needle-holder))
```

Time Needed: Growth of Preservice Science Teachers' Knowledge of Inquiry and Practice of Lesson Design

Augusto Z. Macalalag Jr.
Arcadia University, 450 S. Easton Road, Glenside, PA 19038, macalalaga@arcadia.edu

Abstract: Recent reforms in science education require teachers to develop their notions of scientific inquiry and design effective inquiry-based lessons. This study examined the ways preservice teachers' knowledge of model-based inquiry (MBI) and their ability to use this knowledge in designing lessons developed over time. This study involved 15 preservice teachers enrolled in four consecutive methods courses in the biological certification program. Qualitative procedures were employed to analyze teaching philosophy papers and clinical interviews. Research findings provided evidence to support growth in preservice teachers' knowledge of MBI and its implementation in lesson designs: (a) from teacher-centered and activity-oriented to more student-centered lessons with modeling, and (b) from scripted to more sophisticated modeling practice. This study has the potential to contribute to teacher education research by uncovering the effects of subject-specific methods courses and fieldwork on the development of teacher knowledge and lesson planning practices.

Introduction: The Use of MBI in Preservice Science Teacher Education

Current reforms advocate for science teaching that emphasizes the development of scientific knowledge through engagement in core scientific practices such as modeling, developing explanations, and engaging in argumentation (NRC, 2011). The pedagogical approach associated with teaching science as model building and testing is termed model-based inquiry or MBI (Cartier, Rudolph & Stewart, 2001). MBI in the classroom entails (a) the use of students' prior knowledge to pose problems and generate data, (b) the search for patterns in data, (c) the development of causal models to account for patterns, (d) the use of patterns in data and models to make predictions, (e) the design and conducting of experiments to test models, (f) the revision of models based on evidence, and (g) the conducting of argumentation in light of new evidence (Windschitl & Thompson, 2006). MBI is different from the traditional scientific method approach in that questions are derived from a model that represents observable (e.g., balloon expanding) or unobservable (e.g., collision of molecules inside the balloon) phenomena in the world, rather than being based on what teachers conceive as interesting or doable (Windschitl, 2004).

Teaching through inquiry has a long history in science education. From the early 1960s until today, researchers and educators developed curricula (e.g. Elementary Science Study [ESS] and Biological Science Curriculum Study [BSCS]), standards (e.g., National Science Education Standards and AAAS' Literacy Maps), and professional development programs to help teachers to incorporate inquiry in the classroom (Duschl et al., 2007). MBI has stemmed from research in the philosophy of science that has argued for the central role that models play in scientific inquiry, both in terms of artifacts of scientific thought and as fodder for new scientific explorations. "A scientific model is an abstraction and simplification of a system that make its central features explicit and visible, allowing someone—the inquirer (a scientist, a teacher, or a learner) – to illustrate, generate explanations, or make predictions about natural phenomena" (Harrison & Treagust, 2000, p.2). A model is a set of conceptual understandings that (a) can be used to explain natural phenomena, (b) is continuously assessed and revised in light of new data and evidence, and (c) can be used to make predictions about natural phenomena and thus become a useful guide for future research studies (Cartier et al., 2001).

Theoretical Framework

Lesson planning is a ubiquitous practice for teachers and lesson plans are important artifacts of teaching. The processes of lesson design and the creation of lesson plans are windows into teaching philosophy and strategies. Lesson planning refers to teachers' conceptualization and formulation of courses of action in a lesson, which in turn have a profound influence on teachers' classroom behavior and students' learning (Shavelson, 1987). Research studies have shown the different challenges that inservice and preservice teachers (PTs) faced while engaging in lesson planning and design. Specifically, inservice teachers paid little attention to the scientific theories involved in science lessons (Duschl & Wright, 1989) and were initially uncertain and unaware of the different ways of thinking about concepts among their students (Kazemi & Franke, 2004). However, through professional development, inservice teachers can begin to attend to their students' thinking.

In contrast, PTs have tended to design lessons that were teacher-centered, which did not consider students' prior knowledge and the curriculum (Friedrichsen et al., 2009). In a study that helped PTs increase their knowledge of students, about one-third of the 32 PTs did not consider students' conceptions while designing lessons (Justi & Gilbert, 2002). On the other hand, studies have shown that it was indeed possible for

PTs to conceptualize and design student-centered lessons (Etkina, 2010; Fernandez, 2010). Teaching science through inquiry and developing inquiry lessons specifically around models and modeling is equally challenging for PTs. These challenges include (a) difficulty in letting go of the didactic approach to teaching and moving toward more student-centered instruction (Hayes, 2002) and (b) not referring to scientific theories or models in planning and performing their investigations (Windschitl, 2004). In studies that looked at the ways to improve PTs' knowledge and language of models and modeling, Crawford (2004) was successful in developing PTs' ability to critically think about mechanisms involved in modeling after one semester of engaging in modeling experiences, designing open-ended investigations, and building and testing their own dynamic computer models. However, PTs viewed models differently from the ways that scientists or researchers use models, there was little indication of using modeling in PTs' own teaching, and these PTs failed to use models to design their own investigations (Crawford, 2004; Justi & Gilbert, 2002; Whindschitl & Thompson, 2006). In terms of the teachers' ability to design inquiry-based lessons, Schwarz and Gwekwerere (2007) showed that by using highly-scaffolded frameworks for instructional design, PTs were able to develop lessons that focused on the role of students in the lesson, the progression of students' conceptions in the lesson, and the increased use of different models and modeling to engage students.

In summary, designing and implementing MBI instruction is challenging particularly for PTs. In the Methods courses of my study, PTs engaged in designing, revising, and implementing inquiry based lessons in four consecutive courses. The intensive focus on lesson design was part of Methods II and Methods III, which are described in the next section. Qualitative research methods allowed me to produce comprehensive, in-depth, and holistic descriptions of the growth in PTs' knowledge of MBI and their ability to design lessons through educational philosophy papers and clinical interviews that are meaning-rich (Merriam, 1988). Specifically, the following questions guided my study: (1) In what ways do PTs' knowledge, as demonstrated through course assignments, of MBI develop over the four consecutive methods courses? (2) In what ways do PTs' ability to design model-based inquiry lessons and units change over time? I anticipated that the PTs in my study would increase their attention to students' learning, curricula, and scientific models in their lesson plan and design over the four courses. More specifically, I hypothesized that the initial lesson designs would focus on the teaching models and didactic-approach similar to what Justi and Gilbert (2002) and Hayes (2002) found in their study. Given the findings from the studies by Etkina (2010) and Crawford (2004), I predicted that lessons would begin to focus on MBI with special attention to students' learning and development throughout the lesson. However, compared to the findings of Schwarz and Gwekwerere (2007), I hypothesized that the PTs in my study will consider a more sophisticated version of scientific models in their lesson design given time and emphasis in their methods courses.

Methods: Study Context

My study involved a cohort of 15 PTs (4 male and 11 female) enrolled in a two-year biological science certification program at a large university in the northeast U.S. This graduate program was geared towards two types of students— students that were juniors majoring in the biological sciences or a related field and seeking teacher certification in biological sciences (5-year undergraduate students), and students that had completed an undergraduate degree and were seeking certification (post-baccalaureate students). There were two tracks at this university for science certification—a physical science track to certify physics and chemistry teachers (Etkina, 2010) and a biological science track to certify biology teachers. In both tracks, all PTs had completed at least 30 credit hours in the subject matter (in this case, biology) before entering the teacher education program.

As part of the certification program, the PTs completed four subject-specific methods courses in consecutive semesters (Methods I–IV). In Methods I, PTs engaged in MBI activities, readings, and discourse designed to promote their understanding of scientific inquiry and engender a view of science as theory-building. The goal of the various activities was to provide the PTs with experiences of inquiry from a learner's perspective, and to provide a model of what MBI teaching looks like. In addition to the inquiry activities, in Methods I, PTs also engaged in lesson critique and revisions—an important aspect of teacher preparation (Duncan, Pilitsis, & Piegaro, 2010). A more intensive focus on lesson design was part of Methods II and Methods III. Methods II was essentially a design-based course in which the PTs, in small groups, developed extended model-based inquiry units about selected topics in biology such as photosynthesis, ecosystems, etc. In this course, PTs were introduced to some design frameworks including Learning for Use (Edelson, 2001) and Backwards Design (Wiggins and McTighe, 2005). Lessons and activities in this course were scaffolded to increase PTs' repertoire in analyzing students' prior conceptions and alternative conceptions (Crawford, 2004), decision strategies involved in incorporating models and modeling in lesson design (Schwarz & Gwekwerere, 2007), incorporation of epistemological bases of scientific knowledge in lessons, and experience in teaching inquiry (Windschitl, Thompson, & Braaten, 2008). In Methods III, PTs further developed their abilities to teach inquiry-based lessons and assess students' thinking during their supervised student teaching internship, which lasted 15 weeks. In this course, PTs were given two opportunities to plan, implement, and critically examine extended model-based inquiry lesson sets.

Data Sources and Analysis

In order to capture changes in the 15 PTs' notions of MBI and its implementation on lesson design, I chose data sources that were relevant to issues of lesson design. The data included educational philosophy papers and clinical interviews, each of which is described below.

Semi-Structured Interviews. In Methods I, PTs were asked to address four questions in a teaching philosophy paper: (a) "What are the goals of biology education and what should be taught in high school?" (b) "What are the problems with the current instructional methods?" (c) "What are the best ways to learn and teach science?" and (d) "Describe an ideal lesson in biology." While the questions did not explicitly use the words "scientific inquiry," PTs' answers to the questions gave insights about their implicit notions of science inquiry in the form of investigations or experimentation (not MBI) as well as the characteristics of a good lesson. The philosophy paper was written as a homework assignment and was submitted the second week of Methods I.

The analysis of educational philosophy papers represented a pre-instructional measure of PTs' ideas about the best ways to teach and learn science, as well as the components of an ideal lesson in biology. This baseline point of analysis gave me insights into what PTs may or may not know about MBI and lesson design. The development of my coding schemes proceeded through an iterative process of application to the data set and refinement of the codes to capture relevant emerging themes in the data (Merriam, 1998). I double coded the data in instances when a statement fit into two different categories. I present a complete list of categories and examples of them in the "Findings" section of this document.

Clinical Interviews were conducted with each teacher at the end of each of the four methods courses. The interview protocol had four tasks: (a) defining model-based inquiry, (b) critiquing a lesson, (c) designing a lesson, and (d) evaluating students' written work examples. This paper involved analyses of the third interview task in order to get a sense of PTs' knowledge of model-based inquiry and their ability to develop inquiry-based lessons, respectively. During the third task, PTs were asked to design a lesson based on three objectives given to them. PTs described outlines of a 2-3 day lesson set that would address the provided goals. These design tasks lasted for about 15-20 minutes. Interviews were recorded using audio and video. All interviews were transcribed verbatim.

Before conducting my analyses of the third task (designing a lesson), I blinded (removed names) and rearranged the transcripts from different methods courses to minimize bias. The first coding pass resulted in a list of the different activities in the lesson, such as teachers asking questions to gather students' prior knowledge (naïve models), teachers delivering lectures and demonstrations, teachers voicing students' ideas, hands-on experiments, etc. Through constant comparison of transcripts from interviews at different points in time, I was able to create categories and assign different levels to identify shifts in the nature and quality of teachers' lesson designs with regard to: (1) modeling (level 0—no modeling, level 1—script modeling, level 2A—modeling practice, and level 2B— argumentation) and (2) student-centeredness of lessons (level 0—teacher-centered, level 1—partly student-centered, and level 2—student-centered). After coding all lesson design from the interview transcripts, I identified trends in the categories and subcategories that I mentioned above from the different methods courses. I describe and provide examples of each category and subcategory below.

Findings: PTs' Initial Ideas about Lesson Design

Research questions were addressed by examining the PTs' educational philosophy papers. These papers represent the PTs' initial ideas about the components of an ideal lesson in biology, which I used as a proxy for their ideas of the best ways to teach and learn science. While the assignment did not directly ask for their ideas of MBI as it applies to designing lessons, the educational philosophy papers provided insights into what PTs' thought of as the ideal structure and components of science instruction. The figure 1 below illustrates the salient themes that emerged from my analyses.

With regard to the components of an ideal lesson in biology, the majority of PTs argued for experiential learning through hands-on experiences and experimentation. They argued for the importance of real-life connections to motivate students to learn. A majority of PTs also mentioned lecture-presentation and demonstration as part of the instruction. Several PTs discussed various methods of instruction and allowing students to work independently. These ideas, which reflected a blend of teacher lectures and students' hands-on experiences, seemed to hint, implicitly, at a general view of instruction that involves students, but is heavily directed by teachers. The proposed investigations, discussions, and independent study were used merely to confirm what teachers introduced during lecture presentation or demonstration.

The components of lessons that PTs identified at this stage did not reflect MBI instruction in any way. Only two PTs mentioned the use of models and modeling as part of the lesson. Nora suggested using the model of a cell to review its parts and structure: "Here, we could review the structure of a cell with a model. Looking at the model, the class can locate the nucleus and the chromosomes." However, Nora is using the model to convey information and not as a generative tool to develop students' own ideas. While Jackie, another participant, mentioned incorporating student-generated models in her lesson, she suggested comparing these with other

models or theories in order to provide different perspectives, which somewhat reflected what scientists do as they use alternative models to compare their own:

> After preconceptions were cleared up, the teacher can use inquiry to try to help the student form the concept somewhat by themselves. Then the teacher might want to show the students models and the different theories of the concept to give the students a unique way of looking at the information. (Jackie, Educational Philosophy Paper, Methods I)

Jackie's use of models in her instruction fell short of how scientists use models—to formulate hypotheses from models as well as to test and revise them. The PTs' initial ideas of science instruction invoked hands-on experiences but were heavily directed by teachers, which did not reflect an understanding or valuing of MBI instruction.

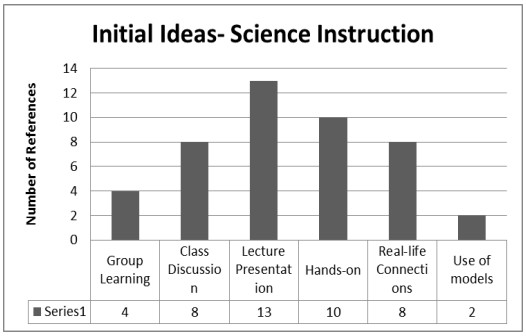


Figure 1. Science instruction based on the teaching philosophy papers.

Model-Based Inquiry

In Methods I, several lessons did not incorporate the modeling process (level 0). These lessons focused on hands-on investigations without eliciting students' prior knowledge. By the end of Methods I and Methods II, PTs had begun to include the language of modeling in their lessons; however, the modeling process was prescriptive and procedure-oriented (level 1). At this level, PTs mentioned the steps of modeling (e.g. gathering naïve models, testing and revising models) but failed to explicitly mention the scientific models that the students were working on. For instance, Patrick said that in the process of MBI, students engage in developing initial models, conducting research and experiments, revising models, and sharing of ideas with other students:

> With the initial model, maybe students would be a little to set to stick with their initial model. They should probably after research, after experiment be ready to develop a new model ... All the while this research and experimentation should be in a group... They are sticking in a group talking to people. They are sharing ideas. And that's it. (Patrick, Clinical Interviews, Methods I)

In this example, Patrick fails to explicitly link the process of modeling to the science concepts that students are working on. His modeling process is generic and prescriptive in that it could be used in any lesson or topic. On the other hand, lessons from interview transcripts in Methods III contained a more sophisticated modeling practice (level 2A) in which target models were explicitly described and connected to the overall modeling practice. In Sean's lesson below, he mentions that after modeling and argumentation, the models should contain and explain the process of photosynthesis:

> After the group model is done, I would have them each group present to the class, post them around the room, and then we would engage in classroom argumentation to see what the differences were and to see if we could reach a class consensus. The models should contain all three of these aspects: how plants get and convert energy—that would be if they put the nutrients chlorophyll and light; equations—inputs and outputs and using glucose as a source of energy... If they explain what the end results are and why they have chosen these results that answer should include— well I included glucose because that is a source of energy. (Sean, Clinical Interviews, Methods III)

Sean's views of models consisted of characteristics and explanations (how plants get and convert energy), which reflected a good understanding of models. However, similar to the lesson sets, only a handful of PTs mentioned argumentation (level 2B) as part of MBI lessons in clinical interviews. These PTs indicated the

use of evidence from experiments to revise and argue for or against a model. For example, Nora's plan was for her students to examine data and use that information to justify their models:

> From that they can analyze that data and make some charts to see the correlation between the different variables in the experiment. After that, I will have them share the results to the class. Every group would share and have the class argue using evidence from their experiments. And have them justify what their argument is about. After argumentation, I will ask them to revise their models. (Nora, Clinical Interview, Methods III)

The majority of lessons did not include argumentation. They did not use evidence to inform their discourse around models. Lessons ended with a revision of models or a lecture. For instance, Catherine mentioned that after investigation, students will revise their models using data from their experiments. Then she mentioned conducting a lecture to summarize the lesson:

> And then I was saying, after that [investigation], have the groups revise the model, then come back and revise the model as a class, using the data. And then, have a benchmark lecture on cellular respiration. (Catherine, Clinical Interview, Methods III)

In addition to lack of argumentation in lessons, there was a decrease in the number of lessons with level 2 modeling (sophisticated modeling practice) and an increase in level 1 modeling (script modeling) in Methods IV. It is not clear why, but this might have been be due to the lower motivation of PTs to complete the interview since this task was their final assignment in the program. Figure 1 shows changes in modeling as part of lessons in clinical interviews.

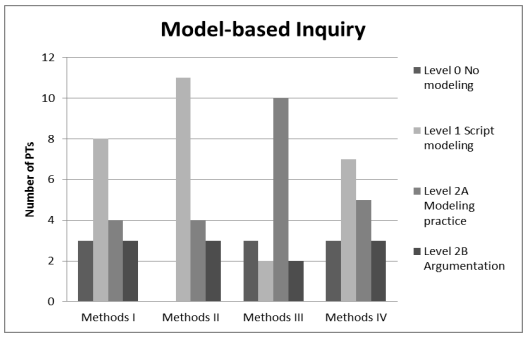


Figure 2. MBI based on clinical interviews.

Designing Student-Centered Lessons

Within the *student-centeredness of lessons* category, PTs described their roles as well as their students' roles in the lesson. PTs' descriptions of the lesson sequence encompassed the different degrees of their consideration of students during instruction: from level 0— teacher-centered (lecture or demonstration) and level 1—partly student-centered (blend of lecture and student investigations), to level 2—student-centered (students' modeling practices) lessons. Lessons that were student-centered tended to have: (a) a decrease in the number of teacher-centered lessons that did not include modeling, (b) an increase in student-centered approaches (e.g. eliciting students' prior knowledge) to learning, and (c) an increase in PTs' ability to anticipate what students knew or would be able to do.

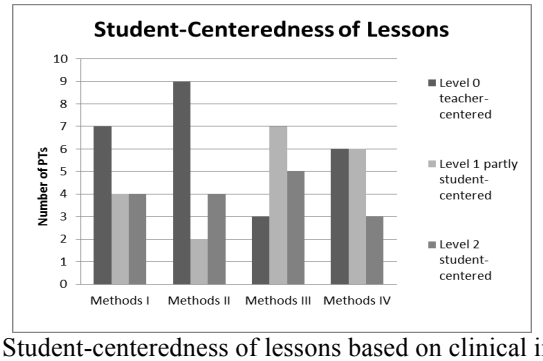


Figure 3. Student-centeredness of lessons based on clinical interviews

Initially, the majority of lessons that the PTs designed in the first two courses were lecture-based and did not consider students' prior conceptions (level 0). At this level, teachers provided information to students in a lecture or demonstration and then asked them to conduct an investigation to confirm what was taught during lecture. As an example, Ava came up with questions for a class discussion, an activity to look at labels in drinks, and a lecture:

> I will have quick question for them: "Where do people get their energy from?" They can look at the ingredients in the labels of bottles of energy drinks and pick two or three ingredients and maybe look for where glucose is coming from. And from there you can let them know that there is a process where glucose and oxygen can give energy and possibly create a naïve model of the process of respiration. And after they are done with more investigation, they can go back and revise their models and present them to class. (Ava, Clinical Interviews, Methods II)

Even when Ava's language included the terms "models," it is noticeable that her lesson was teacher-centered in that the teacher provided the question and explanations to drive the science concepts, and that the lesson did not consider students' models. Moreover, when the interviewer asked her about what she planned to do in between the naïve and revised modeling activities, she answered "that will be a quick lecture [about] the actual respiration process because some of it can be technical." Lessons designed in Methods III and Methods IV, however, had components that were teacher-centered as well as components that included student's hands-on learning experiences (level 1). For instance, Nora described her instruction with the following components: eliciting student ideas, modeling, student investigations, data analysis and presentation, and a lecture to end the lesson:

> I will begin with a guiding question … how can a red wood tree grow so tall? Then I will ask them to create a naïve model to get their prior knowledge regarding the guiding question… I will then introduce an experiment of some sort about photosynthesis, plants, light, CO_2, probes to get CO_2. I will ask them to do an experiment of some kind that will show the changing of concentrations of O_2 and CO_2 by keeping the probes there to see the changes and what is going on. From that they can analyze that data and make some charts to see the correlation between the different variables in the experiment. After that, I will have them share the results to the class... hopefully they are in the same page but if not I will end with a quick benchmark lecture to get everyone in the same page. (Nora, Clinical Interviews, Methods III)

While the lesson had components that were student-driven, the teacher provided the experiments and connections for his/her students. On the other hand, half of the lessons developed in Methods III were student-centered lessons (level 2). PTs' lessons involved active participation of students: learners engaging in scientifically oriented questions, teachers eliciting students' prior knowledge, learners developing or engaging in an investigation, learners gathering evidence and formulating explanation, and learners communicating and justifying explanations. Moreover, these lessons included making connections to scientific knowledge, modeling practices, and/or explicitly mentioned anticipated outcomes from students. Molly described a lesson in which she would elicit and adjust her activity depending on her students' prior knowledge:

> So by starting with plants, you can get them thinking how they make their own food… and how do we get our energy? … I would like to see their prior knowledge …if there is no clear understanding and there is really no prior knowledge then I wouldn't have them model… depending upon their prior knowledge…I know this sounds crazy but I would probably give them some experiments that scientists did to figure out how we get our energy…Like what they ate this for 20 minutes and they were fine or they ate this for 20 minutes and they didn't feel well. So on and so forth. Then I would see if we added something to that prior knowledge… (Molly, Methods III)

Molly's experiment or suggestions are meant to help her students think about the question. Molly described how she would demonstrate an experiment, ask her students to design and conduct their own experiments, and end the lesson by re-visiting the students' initial models:

> I was thinking maybe I would do some sort of small experiment in front of them some sort of to get them going …maybe I would run in place and take my pulse... And then have them design an experiment… I want them to do [their experiments] to see what happens… [And] once they are finish with their experiments, I want us to get together and share what they

did… [next] some sort of small benchmark lesson, just very short, maybe on oxygen, glucose, carbon dioxide and water and how those four work together that may get them thinking… I would have them go back and [and revise] their naïve models… (Molly, Methods III)

Discussion and Implications

Recent calls to refocus science education have emphasized the development of scientific knowledge through model-building and argumentation (Duschl et al., 2007; NRC, 2011). MBI instruction involves an understanding of scientific knowledge as ever-evolving conceptual models of natural phenomena and the scientific practices used to generate, test, and revise those models. However, this type of instruction is difficult to successfully implement, especially for PTs who lack the knowledge, experience, and strategies to teach according to MBI. Specifically, it appears to be challenging for them to develop their own knowledge of scientific models and modeling (Windschitl, 2004), design MBI lessons (Schwarz & Gwekwerere, 2007), and attend to student thinking (Hayes, 2002).

My analysis of the PTs' educational philosophy papers revealed that their initial ideas about science instruction were generally teacher-centered and did not incorporate key aspects of MBI. My findings are similar to those of Hayes (2002) who found that it was challenging for most of his PTs to let go of a didactic approach to teaching and move towards lessons that considered the development of students' own interest. I found, initially, that a typical lesson developed by PTs during the study was mostly teacher-centered. These lessons included teachers asking questions, followed by a lecture, and then guided practice, which ultimately revolved around the lecture material. In most cases, student participation during investigations, discussions, and independent study were merely used to confirm what teachers covered during lectures. Moreover, the majority of PTs in my study emphasized experiential learning through hands-on experiences and argued for the importance of real-life connections and motivation for students to learn. These ideas of science instruction in lesson designs implicitly pointed to the PTs' views of instruction that was heavily directed by teachers. Moreover, similar to the findings of Schwarz and Gwekwerere (2007) who saw that PTs could improve how they think about models but still struggled to incorporate models into their lessons, I found that initially only two PTs in my study mentioned the use of models as part of their instructions. These models were used to convey information, not as generative tools developed from students' ideas, and different from how scientists use models. However, contrary to Schwarz and Gwekwerere (2007), who found that PTs' ideas of models were limited to representing objects or phenomenon, I found that PTs in my study developed a more sophisticated idea about models.

Contrary to the PTs in the studies of Schwarz and Gwekwerere (2007), who attended only one methods course in science, PTs in my study had an opportunity to learn modeling and lesson design in an extended period of time—four consecutive methods courses. However, the successes and struggles of PTs in my study in terms of their knowledge of models, modeling, and lesson design in Methods I (in one semester) were similar to those PTs in studies of Schwarz and Gwekwerere (2007) and Windschitl and Thompson (2006). Specifically, several lessons in Methods I did not incorporate the modeling process. My findings suggest that this kind of learning takes time and that warranted careful scaffolding and multiple learning opportunities for PTs in several methods courses. However, my findings showed a shift from teacher-centered to student-centered lessons after PTs participated in the lesson design and redesign activities, an internship seminar, and reflection assignments in Methods III and Methods IV.

PTs encountered difficulty in embedding argumentation as part of scientific practice. In most cases, PTs ended the inquiry process in their lessons by asking students to revise their models and present them in front of their classmates without any follow up argument or discussion around models and evidence. My findings confirmed what Windschitl et al. (2008) found with their PTs as they engaged, initially, in modeling and argumentation at the beginning of their methods course. Specifically, they also saw that the majority of their PTs mentioned discussing or stating what they learned from their experiments instead of using evidence and models to anchor their arguments.

Teacher-educators can better prepare pre-service science teachers in planning for inquiry-based instruction by providing them with knowledge and experiences of inquiry that focus on science as model-building and -testing. This entails developing PTs' own knowledge of MBI and placing an emphasis on students' active role in scientific practice. My research findings showed that learning takes time, and one methods course in science is insufficient to change PTs' knowledge and practice with regard to lesson design and students' conceptions. Fostering knowledge of MBI and lesson design involves careful scaffolding of activities in consecutive methods courses. Specifically, PTs in my study had a chance to experience MBI as learners and conduct clinical interviews with students to elicit students' conceptions in Methods I, design units and lessons in Methods II and Methods III, conduct action research projects in Methods III, and analyze students' thinking based on data collected during their internship seminar in Methods IV.

References

Cartier, J., Rudolph, J., & Stewart, J. (2001). *The Nature and Structure of Scientific Models*. The National Center for Improving Student Learning and Achievement in Mathematics and Science. Retrieved on December 4, 2010 from http://ncisla.wceruw.org/publications/reports/Models.pdf

Crawford, Barbara (2004). Supporting prospective teachers' conceptions of modeling in science. *International Journal of Science Education, 26*(11), 1379–1401.

Duncan, R.G., Pilitsis, V., & Piegaro, M. (2010). Development of Preservice Teachers' Ability to Critique and Adapt Inquiry-based Instructional Materials. *Journal of Science Teacher Education, 21,* 81-102.

Duschl, R.A., Schweingruber, H., & Shouse, A. (Eds.) (2007). *Taking Science to School: Learning and teaching Science in grades K-8.* Washington DC: The National Academy Press.

Duschl, R.A. & Wright, E. (1989). A case study of high school teachers' decision making models for planning and teaching science. *Journal of Research in Science Teaching, 26*, 467-501.

Edelson, Daniel. (2001). Learning-for-Use: A Framework for the Design of Technology-Supported Inquiry Activities. *Journal of Research in Science Teaching, 38*(3)*,* 355-385.

Etkina, Eugenia. (2010). Pedagogical content knowledge and preparation of high school physics teachers. *Physics Education Research, 6,* 1-25.

Fernandez, Maria Lorelei. (2010). Investigating how and what prospective teachers learn through microteaching lesson study. *Teaching and Teacher Education, 26*, 351-362.

Friedrichsen, P.J., Abell, S.K., Pareja, E.M, Brown, P.L, Lankford, D.M., & Volkmann, M.J. (2009). Does Teaching Experience Matter? Examining Biology Teachers' Prior Knowledge for Teaching in an Alternative Certification Program. *Journal of Research in Science Teaching, 46*(4), 357-383.

Gilbert, J. & Boulter, C. (1998). Learning science through models and modelling. In Fraser, B. and Tobin, K. (Eds.), *International Handbook of Science Education (pp. 53-66).* Dordrecht, The Netherlands: Kluwer Academic.

Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education,* 22(9), 1011-1026.

Hayes, Michael. (2002). Elementary Preservice Teachers' Struggles to Define Inquiry-based Science Teaching. *Journal of Science Teacher Education, 13*(2), 147-165.

Justi, R. & Gilbert, J. (2002). Science teachers' knowledge about and attitudes towards the use of models and modeling in learning science. *International Journal of Science Education, 24(*12), 1273–1292.

Kazemi, E. & Franke, M.L. (2004). Teacher Learning in Mathematics: Using Student Work to Promote Collective Inquiry. *Journal of Mathematics Teacher Education, 7,* 203-235.

Merriam, S. B. (1988). *Case Study Research in Education: A Qualitative Approach.* SF, CA: Jossey-Bass.

National Research Council. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: National Academy Press.

Schwarz, C.V. & Gwekwerere, Y. (2007). Using a Guided Inquiry and Modeling Instructional Framework (EIMA) to Support Preservice K-8 Science Teaching. *Science Education, 91,*158 – 186.

Shavelson, R. J. (1987). Planning. In Dunkin, M. (ed). *The International Encyclopedia of Teaching and Teacher Education (pp. 483-486)*. Oxford, England: Pergamon.

Wiggins, G. & McTighe, J. (2005). *Understanding by Design (2nd ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.

Windschitl, Mark (2004). Folk Theories of ''Inquiry:'' How Preservice Teachers Reproduce the Discourse and Practices of an Atheoretical Scientific Method. *Journal of Research in Science Teaching, 41*(5), 481–512.

Windschitl, M. & Thompson, J. (2006). Transcending Simple Forms of School Science Investigation: The Impact of Preservice Instruction on Teachers' Understandings of Model-Based Inquiry. *American Educational Research Journal, 43*(4), 783-835.

Windschitl, M., Thompson, J., & Braaten, M. (2008). How Novice Teachers Appropriate Epistemic Discourses Around Model-Based Inquiry for Use in Classrooms. *Cognition and Instruction, 26,* 310-378.

Acknowledgments

I owe tremendous gratitude to Dr. Ravit Golan-Duncan for guiding me in conceptualizing and framing this research study. I also would like to thank Moraima Castro, Vicky Pilitsis, John Ruppert and Dr. Nicole Shea for their recommendations and insights on my literature review and analysis of data.

Becoming Agents of Change through Participation in a Teacher-Driven Professional Research Community

Michael Ross, Ben Van Dusen and Valerie Otero, University of Colorado Boulder, 249 UCB, Boulder, CO
Email: michael.j.ross@colorado.edu, benjamin.vandusen@colorado.edu, valerie.otero@colorado.edu

Abstract: This study involves a theory-based teacher professional development model that was created to address two problems. First, dominant modes of science teacher professional development have been inadequate in helping teachers create learning environments that engage students in the practices of science, as called for most recently by the NGSS. Second, there is a lack of teacher presence and voice in the national dialogue on education reform and assessment. In this study, teachers led and participated in a professional community focusing on STEM education research. In this community, teachers became increasingly responsible for designing and enacting learning experiences for themselves and their colleagues. We investigated the characteristics of the science teachers' learning process. Findings suggest that teachers who participated in this model generated knowledge and practices about teaching and learning while simultaneously developing identities and practices as education reform advocates and agents of educational change.

Introduction

The goal of the NSF-funded, Streamline to Mastery project was to develop and investigate a model of science teacher professional development (PD) that is, by design, customized to participating teachers' needs and experiences. Customization is made possible by organizing activity so that teachers increasingly take control of, and responsibility for, their own professional development. Through the task of creating learning experiences for themselves, such as engaging in lesson-sharing and developing their own education research agendas, the nine participating Streamline teachers generated knowledge about teaching and learning. This knowledge took the form of principles of science teaching and learning and propagated in various forms, locally and nationally, to other researchers and practitioners. Workshops for teachers, conference presentations, peer-reviewed publications, and leadership roles in local schools and on national committees have become common means for the Streamline teachers to participate in productive dialogue on science education reform.

Streamline to Mastery was created in response to two problems. The first is the widespread perception that dominant, top-down modes of science teacher PD are largely ineffective (e.g., Borko, 2004). Two of the three authors are former high school physics teachers and come from communities of practitioners that regard top-down PD as costly, disconnected from the needs of teachers, and largely ineffective at instigating meaningful change. Though researchers have theorized about innovative PD experiences (e.g., Grossman, Wineburg, & Woolworth, 2001) and examined factors that appear to make some teacher PD more effective than others (e.g., Borko, 2004; Garet, Porter, Desimone, Birman, & Yoon, 2001), this study seeks to explore a model of teacher learning that deviates significantly from conventional thinking about how changes in practices and broader agency can be realized by science teachers. The second problem that Streamline was designed to address is the conspicuous absence of teachers from the national dialogue on education reform. This dialogue is heavily influenced by groups with little or no connections to actual K-12 classrooms, including policy think tanks, corporate interests, and philanthropic organizations. University researchers make up a large and vocal group of education reformers as well, many of whom are former teachers who conduct research in or about K-12 schools. We argue that even the most well-intentioned of the university researchers are not grounded in the day-to-day reality of being a teacher in the current climate. It is difficult to imagine a conference on best surgical practices being virtually devoid of practicing surgeons, yet that is the unfortunate and unacceptable state of many prestigious conferences in K-12 education research.

In this piece, we present data showing shifts in the roles and practices of the Streamline participants from those typical of practitioners to those associated with teacher leaders and agents of broad educational change. We bring a particular focus to the development and abstraction of general principles of teaching and learning by the Streamline teachers and discuss these practices as they relate to scientific practices and science learning. Finally, we discuss the implications of these findings for the development of future science teacher PD efforts, particularly with regard to the Next Generation Science Standards (NRC, 2013).

Theoretical Framework

Holland, et al. (Holland, D., Lachicotte, W., Skinner, D., Cain, 1998) define identity as a relational phenomenon as much bound to others and all manner of cultural forms as to the person to whom it is attributed. Sociocultural conceptions of identity have been shaped by the work of Vygotsky (1986), which brought the broader social activity into accounts of individual learning. Contemporary researchers have explored relationships between the

constructs of identity, agency, and goals (e.g., Barton & Tan, 2010; Nasir & Hand, 2006; Nasir, 2002) as a means to understand how perceived opportunities and constraints are a medium through which identity and culture mutually and continuously construct one another.

Our work is grounded in the view that identity and cultural practices are mutually constituted. As it relates, we view cultural practices as the aggregate of ever evolving expressions of shared expectations, norms, and values—as expressed through agreed-upon practices—among individuals that are identified by, and identify with, these practices. In this view, each individual navigates and negotiates her own identity development in the world of what is perceived to be possible (or not); meanwhile, this continual negotiation shapes the broader context and thus what is perceived to be possible. Changes in an individual's practices, and in how she positions herself in and through those practices, *constitute* learning.

In this study, we investigate the hypotheses that social contexts in which participation requires authoring of activity (Bahktin, 1993) lead to changes in participation and agency within that social context. For science students, these forms of agency involve capacity and empowerment to engage in scientific practices (Belleau & Otero, 2012; Ross & Otero, 2011; Van Dusen & Otero, 2012). For K-12 science teachers, forms of agency involve capacity and empowerment to instigate change in local science education contexts, such as the teachers' classrooms, as well as the instigation of broader change (Ross & Otero, 2011; Ross, Van Dusen, Sherman, and Otero, 2012). In both settings, we have investigated the hypothesis that learning contexts can be designed to enable learners to become authors of disciplinary knowledge on the basis of collaborative reflection on their experiences. We argue that this agency is associated with identities as scientists (with K-12 students) and instigators of change (with science teachers).

We propose a heuristic of two extremes of K-12 teacher professional development to frame our research: (1) a model in which principles of teaching and learning are provided by experts who administer the PD and are intended to be acquired in some way by the teacher-learner; and (2) a model in which the teacher participants extract principles of teaching and learning collaboratively through evidence-based activities, such as conducting (and reflecting on) their own educational research. For both models, the goal is teacher learning that leads to more effective practice. The design of the Streamline to Mastery PD program is based on the latter model, with the additional goal of the teacher-learners becoming agents of broader educational change.

One intended outcome of learning in the Streamline to Mastery model is that the teachers increasingly engage in the practice of extracting from their experiences generally applicable principles of teaching and learning that may lead to more effective teaching and broad scale educational change. The process of abstracting general principles from reflection on experiences (including systematic observation and social consensus) is known as the process of *induction* (Bacon, 1878). Induction is a complex and iterative form of reasoning involving moving inferentially from specific instances (concrete experience and observations) of a phenomenon toward generally applicable (typically predictive or mechanistic) rules that govern the behavior of that phenomenon. The inductive process in Streamline to Mastery is analogous to the work of communities of scientists that move inferentially from the observation of specific instances of natural phenomenon to the development of general rules about it. The usefulness of these general rules about our interactions with nature, *scientific principles* and *theories,* resides in the predictive and explanatory power they provide. Likewise, communities of teachers may become both practiced in, and come to identify with, the social practice of using evidence to develop and abstract generally applicable principles of teaching and learning. Just as our empirical work investigates the premise that science learners can author scientific principles and models from evidence (Belleau & Otero, 2012; Ross & Otero, 2011) as they increasingly establish identities associated with science, we investigate here the role of authorship in the development of agency and identities as teacher-leaders.

Streamline to Mastery

The Streamline to Mastery program began with four science teachers in 2009, grew to nine teachers in 2010, and ten more are to be recruited in the 2014. Requirements to be in the program included teaching in a high needs school district, completion of a master's degree, and a willingness to share aspects of teaching practices through group collaboration. Additionally, teachers are required to conduct research into their own practices, present at least once per year at a national education conference, and take one graduate level college course (of their choice) per year. The research team, all who participated directly in the program, consisted of the NSF project PI, two doctoral students in physics education research (both of whom were formerly high school physics teachers), and one future physics teacher (who was, at the time, an undergraduate Noyce Fellow). Teachers and researchers met every two weeks to share lessons, plan classroom research, and discuss topics of interest to the teachers. Activities included lesson-sharing, in which teachers and researchers each shared a lesson that they deemed to be effective and "inquiry-oriented," designing and executing education research, and preparing to present at national conferences.

Since its inception, the participants of Streamline have struggled with the apparent lack of structure of the program. From the researchers' perspective, it was difficult to balance the need for structure at the outset in order to establish norms for supporting learning with the undergirding philosophy that the teachers know best

what their own needs are and are capable of learning to design experiences to meet those needs. Thus, the structures, leadership, and mentoring provided by the principle investigator and two graduate students was deliberately tapered over time. As hypothesized, and as the data below show, the teachers increasingly took control of and responsibility for the design and completion of tasks and for the direction, vision, and mission of the group.

Extracting Principles of Teaching and Learning

One example of the inductive process that occurred within the Streamline to Mastery community involved the first cohort of four teachers negotiating the meaning of the term "inquiry." In the first weeks of Streamline to Mastery, the teachers agreed that their prior science teacher preparation and professional development had not adequately prepared them to enact effective science instruction. They were all familiar with the word "inquiry," and knew it was somehow important in science learning, yet, just as a great deal of research has shown (Kang, Orgill, & Crippen, 2008; Wallace & Kang, 2004; Windschitl, 2004), all agreed that their understanding of it was inadequate. The term *scientific inquiry*, which is another term for the scientific inductive process described earlier, is a practice that is critical for the authoring of science ideas in the classroom. Though the researchers were quite aware that the term "inquiry" had become so ambiguous and problematic that it would be abandoned in the Next Generation Science Standards (NRC, 2013), we understood that these teachers felt strongly that they needed to understand how to engage their students in *doing* science in the classroom, regardless of what policy document authors decide to call it. Lesson-sharing was one of the structures used by the teachers to explore the notion of *inquiry* as it relates to science learning in the classroom. Each teacher was asked to bring in their best "inquiry lesson" and share it with the group. The group then discussed the ways in which each lesson was consistent with their current ideas of "inquiry." The researchers studied how the teachers' notions of inquiry changed over time throughout this process. Data collected involved periodic reflections, a scenario-based scientific inquiry survey (Kang et al., 2008), and videos of all of the lesson-sharing and other group discussions (for detailed methods, see (Ross, et al, 2011)). These data were coded using both an *a-priori* coding method based upon the elements of classroom scientific inquiry outlined in *Inquiry and the NSES* (National Research Council, 2000) and an open coding method (Strauss, 1987) that allowed for the emergence of codes and trends in the data.

Through a longitudinal analysis of teacher talk, the researchers constructed a time series representation of teachers' collective negotiation of meaning of the term "inquiry" as inferred from the data. The analysis yielded four phases of meaning negotiation. The four phases are shown in Figure 1, along with a representative sample of the transcript excerpts. In the first phase, the teachers used the term "inquiry" in different and ambiguous ways (I), including "hands on," "real-world," "constructivism," and "best practice." In the second phase, the teachers realized and externalized that they did not have a complete understanding of the term (II).

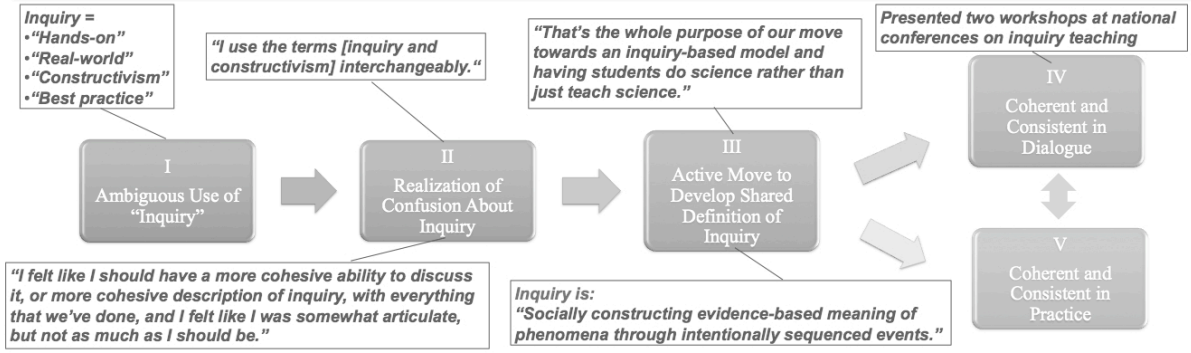


Figure 1. Phases of meaning negotiation of the term "inquiry" among Streamline teachers.

Upon this realization, the teachers decided that they needed to explicitly negotiate better understanding of the meaning and associated practices of inquiry (III). Through concerted effort and communal participation, the group negotiated a shared definition of inquiry as: *"socially constructing evidence-based meaning of phenomena through intentionally sequenced events."* Finally, the teachers used their definition as the foundation for later local and national workshops that they led for other teachers on classroom scientific inquiry (IV). Note that though phase V would be the next logical step in this research, it was not systematically observed or investigated in this study.

The study of the teachers' negotiation of meaning of the term inquiry is an empirical example of the induction of a principle of science teaching and learning. This process was initiated by the teachers and resulted in the extraction of a general principle of teaching and learning. The Streamline teachers moved inferentially from concrete experiences (activity in their own classrooms) and structured observations (lesson-sharing) to a

consensus definition of a central topic of interest in science learning. This example of becoming practiced in induction is just one change that suggests the evolution of the Streamline teachers' identities as expert learners who extract general principles from experience and systematic observation.

Changing Roles, Practices, and Identities

The extraction of principles is one aspect of the learning that was observed among the Streamline to Mastery group. In tandem with the collection of data on the teachers' pursuit of an understanding of "inquiry," group emails and meeting videos were systematically examined for indications of changes in the nature of the teachers' participation as they took on new roles and responsibilities associated with the program. These changes in practice were taken as indicators of increasing agency among the teacher participants. The group emails were analyzed to determine from whom they were sent (teacher or researcher), the date of their origination, what their primary topic was, and if they represented a new line of discussion or were a reply to another email. Transcripts of the meeting videos were coded for the introduction of new agenda items and who introduced them (for detailed methods, see (Van Dusen & Otero, 2012)).

As is shown in Figure 2, the researchers sent the majority of the emails in 2009, but over time the teachers began to send a larger share of the total number of emails. The total number of emails in a given month ranged from eight to fifty four. During period (a) in Figure 2, the majority of the emails were from researchers and focused on scheduling meetings with the occasional email about meeting agenda items. During period (b), the teachers' email volume exceeded those of the researchers. During period, (c), teachers sent an increased percentage of the emails in April. In May, researchers' emails exceeded those of the teachers, largely because the researchers were providing guidance and feedback to teachers as they prepared for their first national conference.

The same set of emails was analyzed to determine whether the teachers or researchers were initiating the *conversation threads*. In this analysis a slightly different pattern emerges. Figure 3 shows the percentage of new email conversation threads by month, again broken down into three time periods. During period (a), researchers initiated nearly all of the new conversations. During period (b), threads initiated by teachers and researchers were evenly balanced with the exception of the March. In March, the teachers were preparing presentations for a regional conference, which required significant communication among the teachers. During this time the researchers primarily acted as resources in answering teacher questions. During period (c), email conversations begun by teachers and researchers were nearly balanced. Figure 2 suggests that teacher participation increased over time, and Figure 3 suggests that the teachers increased their involvement and initiative for leadership.

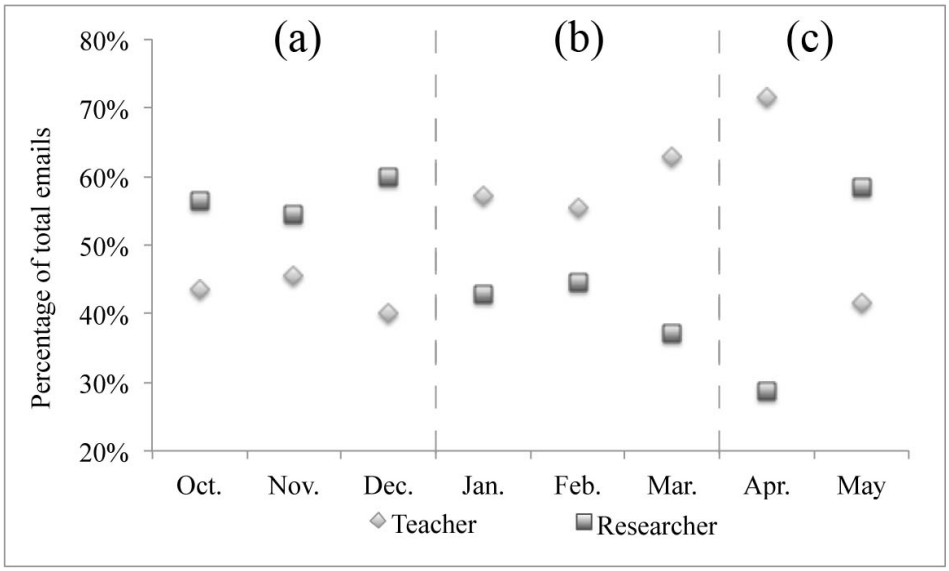


Figure 2. Percentage of total emails sent each month.

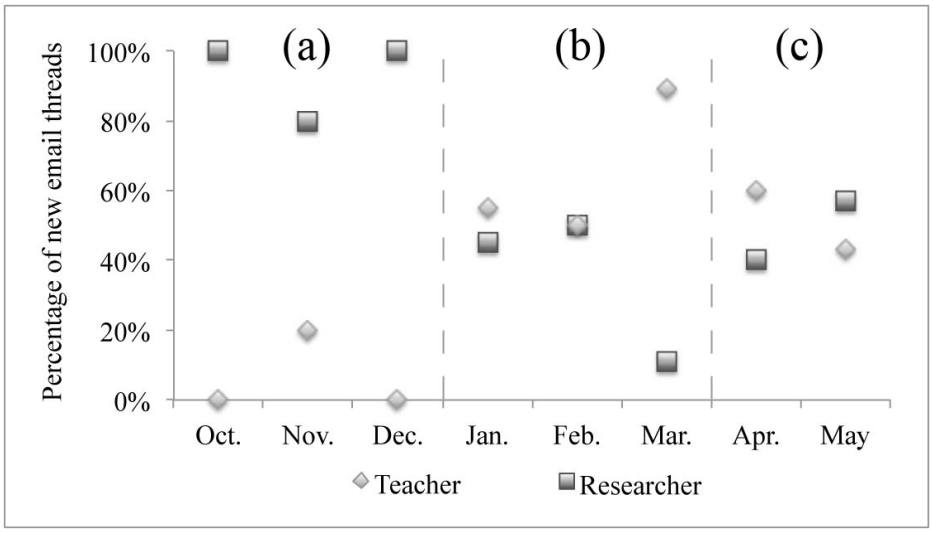


Figure 3. Percent of new email discussion threads.

Videos of group meetings were also analyzed to explore possible changes in who, teachers or researchers, set agendas. We averaged data within each time period (early, middle, and late). As shown in Figure 4(a), researchers provided all of the meeting agenda items during the early time period. The middle time period shows substantial growth in the percentage of agenda items provided by teachers. During the late section (c), the teachers provided the majority of the agenda items. The change that takes place between (a) and (b) came largely from the teachers beginning to take charge of the meetings, ultimately taking on a majority of the agenda setting responsibility.

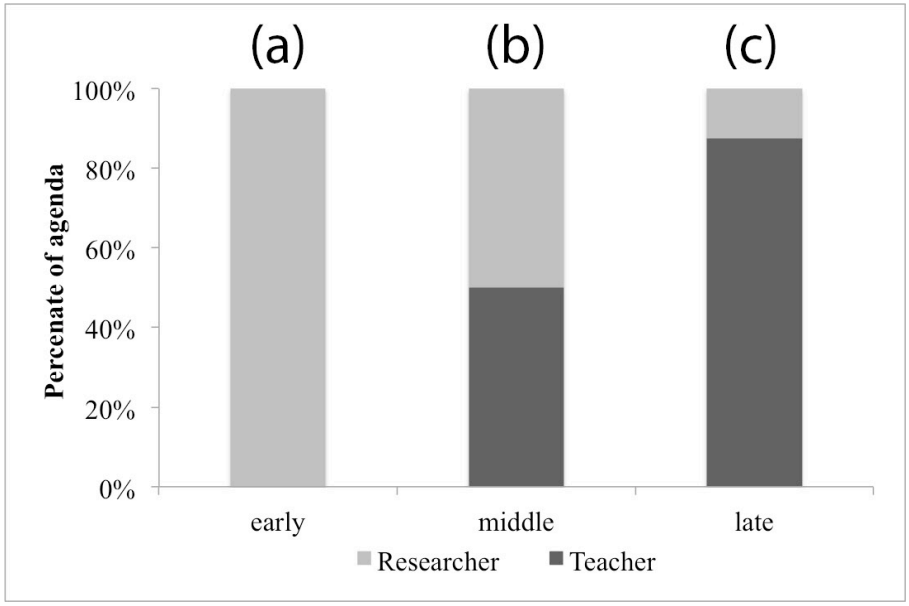


Figure 4. Agenda items from teachers and researchers.

More recently, we examined the type and frequency of the participating teachers' practices of propagating knowledge generated by the Streamline community to other researchers and practitioners. The teachers are responsible for updating a communal spreadsheet with any formal presentations of research findings and teaching practices that they give. This spreadsheet serves as a way for the teachers to keep informed of the work that their peers have done and gives the researchers an easy way to track the community's activities.

As is shown in Figure 5, the instances of sharing knowledge generated within the Streamline community has increased over the four-year period since the group was formally created. Prior to the formation of Streamline, the teachers did not have any role in formally sharing their knowledge of teaching and learning. As a result of being members of the Streamline community, each teachers' formal sharing of knowledge about teaching and learning increased from zero presentations in 2010 to 27 presentations this year (2013), with several more presentations planned before the year ends. As the teachers have taken on stronger teacher-

researcher identities, they have also taken up the practice of prioritizing the generation of new knowledge and sharing their learning with others. Not only is the propagation of research findings a practice associated with teacher-leaders and agents of educational change, but the teachers' activity of executing research projects constitute both learning and becoming. Through this research, the teachers generated new knowledge about the topics of study *and* took up identities of those who are competent at learning through rigorous research.

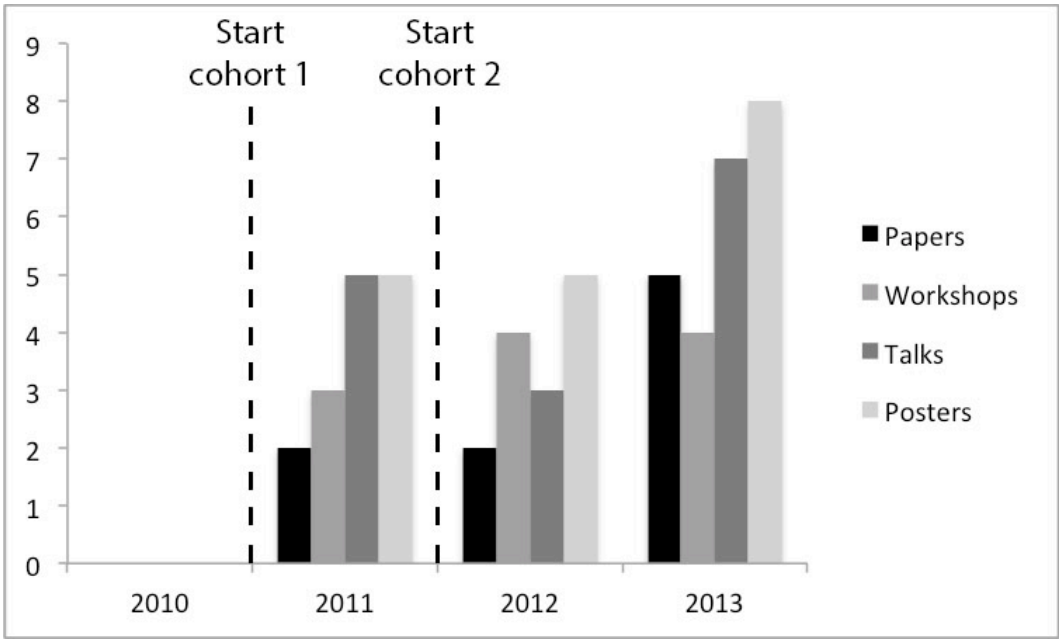


Figure 5. Propagation of Streamline learning to others.

The propagation of knowledge generated through research is tangible evidence of the teacher participants' engagement in the process of induction. This research resulted in the extraction of principles of teaching and learning science. Just as the first cohort of Streamline teachers negotiated meaning of the term "inquiry" in the early stages of the program, all nine participants have engaged in an analogous, and ostensibly more rigorous and systematic, processes of extracting principles of teaching and learning science. Furthermore, many of these findings were recognized as legitimate contributions to the body of education research through peer-review. Moreover, it is the authors' view that the learning and becoming through the disciplinary practices exhibited here may support continued agency in the realms of local and national education policy. The processes of extracting and sharing principles of teaching and learning and of supporting claims about effective policies for education reform with peer-reviewed, evidence-based research has afforded some Streamline to Mastery participants unanticipated levels of agency in national reform efforts. One exemplar of this agency is the participation of a Streamline teacher in the American Chemical Society's (ACS) efforts to launch a chemistry teacher organization comparable to the American Association of Physics Teachers (AAPT). This teacher played a key role in persuading the ACS, through the presentation of her research and experiences with Streamline to Mastery and AAPT, to fund and institutionalize the American Association of Chemistry Teachers, an organization dedicated to the continued development and support of effective chemistry teaching at all levels.

Conclusions and Implications

In this paper, we have examined the learning of teachers participating in a teacher-driven professional development program. Through a sociocultural perspective of identity development, these data demonstrate that Streamline teachers became practiced in various ways that suggest the development of identities as evidence-based advocates of educational change. By extracting their own principles of teaching and learning through lesson-sharing, reflection, and research, the teachers participated in communities that identify with the practices of using evidence to inform and advocate for science education reform. The abstraction of principles of teaching and learning and the taking on of identities of evidence-based education reformers is indeed an example of *learning and becoming in practice.*

It is no coincidence that the model of teacher learning that we espouse is modeled after the central practice of science, *induction.* Just as the core practice of science is to infer from evidence and abstract through consensus the principles, theories, and models that predict, explain, and fundamentally change the ways that humans understand and interact with the natural world, we see generality in this practice. This process of generating new knowledge is applicable to teachers' generation of knowledge as well as formal and informal K-

16 science learning. In our interpretations and discussions of data on learner-driven models of PD and science learning, we must point out what we see as the critical role of the fuller participants (Lave & Wenger, 1991) in these activities. In our view, the development of principles in science, science learning, and teaching and learning is guided heavily by the expectations, norms, and pedagogical expertise of the more knowledgeable others of a community. Just as effective science teachers must be able to engage students in the practice of scientific induction, administrators of PD must strike some productive balance between providing structure and guidance to the learners and providing space and opportunities for the expressions of agency and the identity development valued by the designers and participants in these contexts.

The body of research we are generating suggests that authority-based models of learning science and of learning the practices of effective science teaching may never realize the agency and identity development that is associated with learning in which the making of inferences from evidence and the abstraction through consensus and subsequent reification of these inferences are the domain of the learners. We must continue to explore the premise that for meaningful learning to occur, the very development of abstract, general principles must be performed by the learners, as opposed to delivered to them by some authority, whether it be text, teacher, or some digital medium. Of course, these broad claims warrant further study, but we assert that the learning demonstrated by the data presented about Streamline to Mastery supports the assertion that learner-driven models of induction can be associated with marked changes in agency in valued practices. As the Next Generation Science Standards (NRC, 2013) impose a new set of demands on current and future science teachers, we can be reasonably sure that the dominant, top-down models of teacher (and student) learning will be inadequate to the task of supporting teachers in enacting these standards. Learning through a general model of learner-driven, evidence-based induction has the potential to clarify the central goals of science education, science teacher education, and to more accurately represent science in our classrooms.

References

Bacon, F. (1878). *Novum Organum*. Oxford: Clarendon Press.

Bahktin, M. (1993). *Toward a Philosophy of the Act*. Austin, TX: University of Texas Press.

Barton, A., & Tan, E. (2010). We be burnin'! Agency, identity, and science learning. *The Journal of the Learning Sciences*, *19*(2), 187–229.

Belleau, S., Otero, V. (2012). Critical Classroom Structures for Empowering Students to Participate in Science.
In S. Rabelo, C. Singh, & P. Engelhardt (Eds.) *2012 Physics Education Research Conference Proceedings.* Melville, NY: AIP Press.

Borko, H. (2004). Professional Development and Teacher Learning: Mapping the Terrain. *Educational Researcher*, *33*(8), 3–15. doi:10.3102/0013189X033008003

Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, *38*(3), 181–199. doi:10.3102/0013189X08331140

Garet, M. S., Porter, a. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What Makes Professional Development Effective? Results From a National Sample of Teachers. *American Educational Research Journal*, *38*(4), 915–945. doi:10.3102/00028312038004915

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a Theory of Teacher Community. *Teachers College Record*, 942–1012.

Holland, D., Lachicotte, W., Skinner, D., Cain, C. (1998). *Identity and Agency in Cultural Worlds*. Cambridge, MA: Harvard University Press.

Kang, N.-H., Orgill, M., & Crippen, K. J. (2008). Understanding Teachers' Conceptions of Classroom Inquiry With a Teaching Scenario Survey Instrument. *Journal of Science Teacher Education*, *19*(4), 337–354. doi:10.1007/s10972-008-9097-4

Lave, J., Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Nasir, N. S., & Hand, V. M. (2006). Exploring Sociocultural Perspectives on Race, Culture, and Learning. *Review of Educational Research*, *76*(4), 449–475. doi:10.3102/00346543076004449

Nasir, Na'ilah Suad. (2002). Identity, Goals, and Learning: Mathematics in Cultural Practice. *Mathematical Thinking and Learning*, *4*(2-3), 213–247. doi:10.1207/S15327833MTL04023_6

National Research Council. (2013). *Next Generation Science Standards: For States, By States*. Washington
DC: The National Academies Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. *Science Education*. Washington D.C.: National Academies Press.

Ross, M. & Otero, V. (2012). Challenging Traditional Assumptions of High School Science through the PET Curriculum, in S. Rabelo, C. Singh, & P. Engelhardt (Eds.) *2012 Physics Education Research Conference Proceedings.* Melville, NY: AIP Press.

Ross, M., Van Dusen, B., Sherman, S., & Otero, V. (2011). Teacher-driven professional development and the

pursuit of a sophisticated understanding of inquiry. In Rebello, N. S., Engelhardt, P. V., & Singh, C. (Ed.), *Physics Education Research Conference* (pp. 327–330). doi:10.1063/1.3680061

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge Univ Press

Van Dusen, B., Otero, V. (2012). Influencing Student Relationships With Physics Through Culturally Relevant Tools. In P. E. S. Rabello, C. Singh (Ed.), *Physics Education Research Conference Proceedings*. Melville, NY: AIP Press. Retrieved from http://arxiv.org/abs/1209.1365

Wallace, C., & Kang, N. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets. *Journal of Research in Science Teaching*, *41*(9), 936–960. doi:10.1002/tea.20032

Windschitl, M. (2004). Folk theories of inquiry: How preservice teachers reproduce the discourse and practices of an atheoretical scientific method. *Journal of Research in Science Teaching*, *41*(5), 481–512. doi:10.1002/tea.20010

‘Mangling’ Science Instruction: Creating Resistances to Support the Development of Practices and Content Knowledge

Eve Manz, University of Colorado at Boulder, School of Education, UCB 249, Boulder, CO 80309
eve.manz@colorado.edu

Abstract: This paper explores Pickering’s “mangle of practice” as a tool for designing classroom environments that integrate content knowledge and scientific practices. I describe the design of science instruction for an elementary school class, characterizing how it built from “The Mangle.” I then identify two forms of activity that emerged, *defining attributes* and *mapping between the experiment and target system.* I show how each became a useful practice as material resistances in the system were made public and describe how each served as a site in which concepts and practices were evaluated in relation to each other. Finally, I comment on implications for the design of learning environments that make knowledge-building practices both accessible and relevant to students.

Introduction

There is a consensus that science learning environments should integrate content knowledge and scientific practices so that students learn to generate, use, and support scientific ideas (National Research Council, 2012). In this paper, I explore a concept from the Science and Technology Studies literature, Pickering’s (1995) notion of “the mangle of practice,” as a tool for designing activity that both establishes a need for scientific practices and provides a context for developing content knowledge. I share how Pickering’s ideas guided the design and analysis of a plant growth experiment conducted in a third grade classroom.

The Mangle of Practice

Pickering’s exploration of “The Mangle” elaborates how, in professional activity, scientific practices and ideas become needed, are made problematic, and are revised in light of each other. Pickering conceptualizes science as a dance of human and material agency comprising iterations of resistance and accommodation. Scientists enact their agency by developing hypotheses, procedures, machines, and measures, which they apply to material phenomena. The world responds by doing something, generally something unexpected and somewhat mysterious; it *resists* its capture by human agency. Scientists then must engage in *accommodation,* developing new goals, practices, and understandings. On this view, practices and understandings are tuned and stabilized in relation to each other. When experiments do not perform as expected, scientists reconsider both their material procedures (e.g. experiments or measures) and their conceptual accounts, that is, their understanding of the phenomenon and how the experiment represents it. Producing a scientific finding involves making procedures, conceptual accounts, and results hang together. Therefore, material puzzles are essential to the development of both practices and concepts: they destabilize them, establishing a need to reconsider each in light of the other. These processes are evident in historical analyses of scientific activity (Gooding, 1990) and ethnographic accounts of laboratories (Nersessian, 2012).

Why the Mangle Might Be Useful in Classroom Settings

Scientific practices do not transfer unproblematically from expert settings into classrooms; understandably, their purposes and forms tend to be unfamiliar to students (Hogan & Corey, 2001). Two prominent instructional strategies for introducing scientific practices have been making their structures explicit and simplifying the demands of applying them. However, it is becoming clear that students can adopt taught forms without understanding their purposes or finding them meaningful for their activity (Berland & Reiser, 2011; Kuhn & Pease, 2008). In response, researchers increasingly seek to design contexts that establish a need for practices and to study their development over time. These approaches are consistent with sociocultural accounts, which emphasize that practices are constituted in community activity as members seek to align behavior and accomplish goals (e.g., Wenger, 1998). They involve a shift in how we frame “scientific practices” in classrooms. Rather than viewing them as forms of activity in which scientists engage and that we seek to introduce to students, we might define something as a “scientific practice for students” if it is constituted by a classroom community for a function that is important in their scientific activity.

The Mangle provides a framework for considering both when students might experience a need for scientific practices and what it might mean to adapt those practices in extended activity. In Pickering’s account, practices emerge and are refined in order to cope with resistances; that is, they are made necessary by the material and uncertain nature of scientific activity. There is some evidence that purposefully designing materiality and uncertainty into learning environments can situate the development of sophisticated scientific processes (Lehrer, Schauble, & Lucas, 2008; Roth & Roychoudhury, 1993). However, to date, there have been

few accounts of how materiality and uncertainty are made visible in instruction, how they situate new forms of activity, and how those forms of activity are constituted in classrooms as practices with identifiable functions.

In addition, the Mangle explicitly integrates conceptual work into descriptions of scientific practices, providing a lens for considering the development and use of content knowledge. Pickering's description is consistent with recent accounts that frame ideas as resources for navigating activity, rather than as units of declarative knowledge (Hall & Greeno, 2008). These resources might include ways of attending to significant aspects of situations, organizing information, and making inferences. Applying the Mangle to classroom learning environments supports an important shift from equating content knowledge with the explanation that is the *target* of an investigation toward fine-grained consideration of the ideas that students draw on to navigate their work *throughout* the investigation.

In the remainder of the paper, I apply the Mangle to explore the following questions: (1) How can we create resistances in learning environments that destabilize practices and ideas, creating a need for students to consider and tune the two in relation to each other? (2) What does it look like for students to engage in this process? I describe the design of an investigation conducted with an elementary school class, characterizing how it built from Pickering's ideas. I then identify two forms of activity that emerged in the classroom, *defining attributes* and *mapping between the experiment and target system*. I show how each practice became useful as resistances in the experiment were made public and how each involved tuning concepts and practices in relation to each other. Finally, I comment on implications for the design of learning environments that make knowledge-building practices both accessible and relevant to students.

Design

The context of this work was a multi-year design study conducted with third-grade students (ages 8 & 9) in an urban school (approximately 70% free and reduced lunch). The students' teacher had 30 years teaching experience and had participated in four years of professional development around modeling-based science instruction. We engaged students in developing explanations of "the wild backyard," a trapezoidal-shaped area behind their school (1). The school wall cast a changing pattern of shade on the backyard, resulting in differential sunlight and moisture and related patterns of plant distribution. The target explanation was one of *differential success:* different plants are successful in different amounts of sunlight. This explanation is initially very challenging for students to construct, as they find it difficult to privilege and relate light and plant presence among the myriad potential variables in the backyard (Manz, 2012).

I report here on one phase of the second year of the design study, conducted with a class of eighteen students (13 male, 5 female). The "plant growth experiment" was conducted between the end of February and the beginning of May. By the start of this phase, students had begun to identify "sunny" and "shady" areas of the wild backyard as well as areas that they thought received "some sun and some shade." However, they were confused about the effects of light, partly due to the fact that many of the plants they had been studying in the fall had died in the areas where they had been located (this was due to seasonal change and life cycle processes, but it was a puzzling result for students). We introduced the Wisconsin Fast Plant™, which completes its life cycle in seven weeks, as a context for exploring both the effects of light and plant life cycle processes.

We designed the plant growth experiment to engage students in the mangle of practice as they developed explanations of differential success (Figure 1).

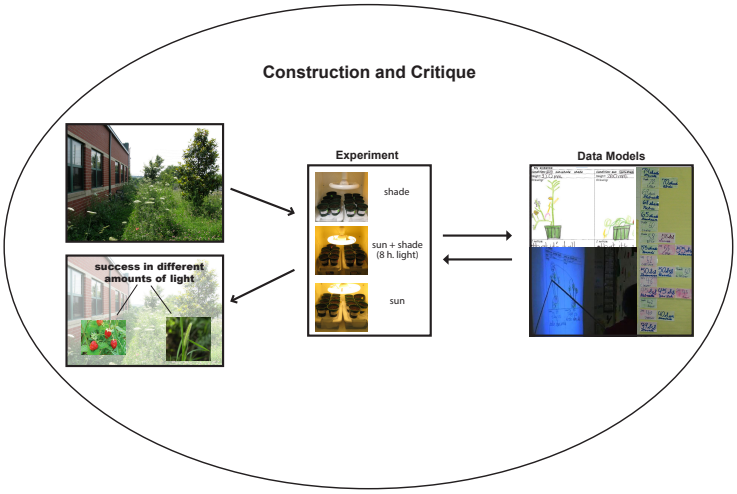


Figure 1. Design

 © ISLS

We introduced a material model of the backyard in the form of an experiment in which students placed Fast Plants in different conditions to mimic those that they had identified in the backyard setting ("sun," "shade," and "sun & shade," referring to areas that were sometimes in shadow). As indicated by the arrows in Figure 1, the processes of designing the experiment and applying its results to develop explanations of the backyard involved significant uncertainty, and therefore constituted sites for experiencing the Mangle. Students had to grapple with how the experiment represented the conditions in the backyard and how its results informed their understanding of plant needs in the backyard (How should they represent light? Did the Fast Plants represent all the backyard plants?). In addition, rather than telling students what about the plants might be important to observe and how to observe and record it (e.g. directing them to graph plant height), we conceptualized the development of data models as another site for mangles to emerge. Numerous plant attributes might be important to observe and compare; these attributes changed over time and often contradicted each other. Through the design choices above and their implementation in the classroom, we sought to position students as constructing and critiquing the system portrayed in Figure 1 (Ford, 2008; Gresalfi, Martin, Hand, & Greeno, 2009). Forms of activity in which we engaged students included small and whole group discussions about how to set up and interpret the experiment, individual writing in science journals, and class "research meetings" in which different students presented ideas about which plants were more successful and took questions from their classmates.

Methods

Consistent with methods for design-based research, conjectures about students' practice and productive means to support development were iteratively developed and refined over the course of the study (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). As the lead researcher, I worked with the teacher and larger research team to design all activity, reviewed evidence of student learning to support ongoing re-design, and was as an active participant, sometimes a co-teacher, during lessons. Data sources included video-recordings, field notes, student work, classroom artifacts, and interviews. During each lesson (n=16, 1-1.5 hours each), a video was made of whole group discussion. During individual and small group work periods, one camera followed the teacher, while one to two additional cameras were used to capture the work of groups.

Retrospective analysis of the data focused on describing normative, or "taken-as-shared," practices (Cobb, Stephan, McClain, & Gravemeijer, 2001) and understanding the purposes they served for students. That is, I sought to develop a description of what counted as practices *in this classroom community,* rather than first specifying and describing a desired practice, then looking for evidence that students were developing aspects of it. However, I was also guided by disciplinary considerations, in that I focused on interactions around the experiment, data models, and explanations of the backyard, as illustrated in Figure 1.

I began by using grounded analysis (Glaser & Strauss, 1967) to describe students' participation in construction and critique. I was interested in which aspects of the system described in Figure 1 were framed as the target of claim-making, justification, and disagreement, rather than as subject to recall or review. For example, students developed and argued about which plants were more successful and how to measure the plants, while the number of days that the plants had been growing was routinely treated as unproblematic and subject to recall or reference. I then sought to understand which forms of construction and critique became practices for students, in that they were repeated, involve broad participation across the class, were initiated by students as well as teachers, and appeared to serve fairly stable (though not always identical) purposes.

After developing a set of categories to describe practices, I divided the data set into activity phases (e.g., a discussion of how to measure the plants). For each activity phase, I asked what practices students were engaged in and described how ideas about plants were used. I also looked for and described evidence of *accommodation,* in that students positioned practices or ideas as problematic or needing elaboration. Finally, I made conjectures about why students were using a practice or idea, with an eye toward noting any resistances they were grappling with.

In this paper, I focus on two forms of activity that developed into repeated classroom practices, *defining attributes* and *mapping between the experiment and target system.* These practices were chosen because there were multiple instances of each and each showed evidence of accommodation in reaction to system resistances, but they emerged at different times in the investigation and appeared to serve different purposes, providing an interesting contrast. I then conducted a more detailed analysis of these two practices. I identified each instance of the practice and bounded it within an episode in which it was initiated and used, resulting in 48 episodes (32 for defining attributes and 16 for mapping). For each episode, I described who initiated it; asked whether there was evidence of student construction, critique, and accommodation; and analyzed how ideas about plants were brought to bear on activity. Examining patterns across episodes allowed me to develop a description of how each practice emerged and was appropriated, what role resistances played, and how ideas about plants were brought to bear on use and accommodation of the practice.

Findings

In this section, I share my analysis of the two focal practices, *defining attributes* and *mapping between the experiment and target system*. For each practice, I present a brief description, then address two questions: (a) How did the practice emerge? (b) How did it help students consider and develop ideas?

Defining Attributes to Cope with Changing Plants

As students looked at plants and discussed which condition the plants were most successful in, they generated and observed many plant attributes to support their claims, including "big," "height," "growing," "light green," or "dead." They engaged in *defining attributes* when they described attributes in more specific terms that allowed others to see and compare them across plants and when they requested that an individual or the class construct a more specific description. For example, early in the investigation, Ellen noted that one of her plants had what she called "a bump." Azhad initiated an episode of definition by asking "Is that bump, is it part of the leaf or part of the stem?" prompting a series of conjectural definitions that named the bump as a precursor to another feature, such as a leaf. There were 32 episodes in which students requested and/or proposed definitions.

Emergence of the Practice

Teachers (2) modeled and asked for definitions across the duration of the investigation. Most of the teacher-initiated definitional episodes began when a student used a term such as "growing" or "big" to compare plants and a teacher asked for elaboration of the chosen attribute. For example, when Charles noted, "the sun and shade is smaller than the shade," Mrs. W. asked him what he meant by "smaller," then continuing to press him until he had defined *size*, which could encompass a variety of attributes, as *height*. Across the data corpus, Mrs. W. and I initiated sixteen episodes of defining attributes. Eleven of these had a similar structure to the episode above, in that we followed a students' use of an attribute by asking "What do you mean," "How do you know," or "What tells you" and students responded by elaborating with more specific descriptors.

Students initiated half (n=16) of all definitional episodes. Many were attempts to identify attributes in the face of plant change, which constituted a resistance for students. For example, when Dante claimed that the plants in the sun & shade condition were doing best because the plants in the sun condition were dying, Azhad disagreed, beginning an episode in which definitions were proposed and challenged. (3)

1. Azhad: No, because I don't see no one dying.
2. Dante: You don't see those leaves that are getting dried up? I know that some [plants]
3. Brady: [How] do you know it's dead though?
4. Britney: (*undecipherable*) drying up
5. Dante: [I know it's]
6. Alex: [Those are] OLD leaves.
7. Jasmine: [Those are] the seed leaves, [that's why they're dying.]
8. Chad: [No they're not.] (*walks over to the lightbox)*
9. Azhad: Those are the [leaves that grew first.]
10. Alex: [Those are the seed leaves.]
11. Madison: [Those are the old leaves.] They're trying to grow new ones.

This excerpt exemplifies how definition emerged as students struggled to see the same thing in the face of a changing system that resisted description. Both Azhad and Brady problematized the notion that death could unproblematically be "seen;" Azhad when he argued that he did not see any dying plants (Line 1) and Brady by positioning death as an inference that needed to be justified (Line 3, "How do you know it's dead though?"). In response, Dante defined dying by bringing in a new, more specific attribute, "leaves that are getting dried up." In turn, students contested this definition. They argued that the leaves drying up were the "old leaves," or seed leaves that they had learned come first and provide the initial food to the plant, and that their drying up might not have anything to do with death. After Mrs. W. reviewed students' characterization of the leaves as seed leaves, Dante went back to the plant boxes and said, "No, I see some spiky leaves that are brown," referring to the true leaves that come later in a plant's life. Across the episode, Dante engaged in accommodation, progressively refining his definition so that others could see the plants in the sun condition as dying and, he hoped, agree that they were not getting what they needed. He needed to do so to contrast the attribute he sought to apply to the plants, death, to the normal processes of maturation claimed by other students.

Across the data corpus, twelve out of the sixteen episodes of definition initiated by students involved struggling with how the plants were changing over time. The students used similar constructions as teachers, in that they asked "How do you know" and "What do you mean," but they applied these constructions

to a different subject (change over time) and were more likely to use them when engaged in disagreement, as in the episode above. Therefore, it appeared that definition was a practice that they found useful for their own purposes, which involved developing shared ways of seeing plant features in the face of life cycle changes.

Definition as a Site for Conceptual Work

Definitional episodes were rich sites for the recruitment and refinement of the forms of ecological thinking that we sought to develop. Across these episodes, ideas about plants were *differentiated*, *related*, and called on as *mechanisms or predictable processes* to support claims. Consider the disagreement about Dante's claim. As Dante was challenged by his classmates to show that the plants were dying, and conversation shifted to definition, he brought in new aspects of the system (the plants' leaves, Line 2), *relating* them to death. When his classmates, in turn, contested the notion that brown leaves indicated death (Lines 6-11), they did so by proposing an alternative *mechanism,* in this case a *predictable process* of maturation, to account for leaf change, arguing that the old leaves that were trying to dry up so they could grow new ones. As students contested the definition of death, leaves were *differentiated* into seed leaves (or "old leaves") and true leaves ("spiky leaves"). Here, definition was a highly conceptual process that pitted plant maturation against death.

Over the course of their work, students appeared to develop stable accommodations, in that they increasingly defined attributes in relation to life cycle processes. For example, on April 21, Brady indicated that his plant was successful because it had "buds where flowers will grow." Here, the attribute of the "bud" was identified and defined in terms of a future feature. In fact, as we asked students to conclude which plants were more successful, this predilection caused difficulty for the classroom teacher, who was ready to end the investigation and decide that the plants in the sun condition were more successful because they had produced seedpods. While students privileged seedpods as a sign of reproduction and therefore success, they disagreed that the sun plants were more successful and supported their counterclaims with prolonged argumentation about what counted as a seedpod. Several students argued that the pistils on the sun & shade plants, where flowers had fallen off but no seeds were growing, were "newborn" seedpods where seeds would grow. Steven interrupted a count of seedpods, saying "There's this question I wanted to ask people, what if their seedpods are dead, does that count as a seedpod?" In these episodes, students framed the seedpod as a maturing, dying entity, complicating its definition, which they considered necessary for a shared understanding of which plants were more successful.

These definitional episodes showed evidence that a resistance (i.e., the changing nature of the plants) destabilized, and supported the development of, both practices (identifying and defining plant attributes) and conceptual accounts (maturation, death, and reproduction; major life cycle concepts). In order to contest and develop definitions, students needed to call on ideas about plant growth. Therefore, definition was a context within which these ideas were useful and became the subject of argumentation. In this way, definitional practices and understanding of plant life cycles were tuned in relation to each other, as Pickering describes.

Mapping between the Experiment and Target System to Explore Differing Results

One way in which this experiment differed from many investigations conducted with elementary school students is that it was explicitly designed to model another phenomenon that students experienced: the backyard system. At several points, students *mapped between the experiment and target system*: they thought about the ways that the experiment was and was not like the backyard and the consequences differences might have. Similarly to *defining attributes,* students engaged in this form of activity when requested to do so by the teacher, but also initiated episodes, in this case by proposing important similarities or differences to consider and challenging the mappings that others made between the systems. Sixteen episodes were located in the data set.

Emergence of the Practice

Throughout the investigation, the teacher asked students to make explicit mappings between the backyard and the experiment. For example, as students were setting up the experiment, she asked them whether it was OK that the shade condition (the lightbox with the light off) let some light in, as it was translucent. Students decided to block the back of the lightbox with cardboard to be like the school wall that blocked light, but argued that it was fine that the sides let some light in, because light could get into the shady areas of the backyard from the sides too. On seven occasions, the teacher explicitly asked students to make mappings, either by focusing on making connections between the conditions of the experiment and the backyard or by asking students to use the experiment to make predictions about where Fast Plants would be successful in the backyard.

During several of the conversations seeded by the teacher, students initiated the discussion of aspects that did not map and, on a few occasions, spontaneously discussed the implications of these misfits. Consider, for example, the conversation below.

1. Mrs. W: How does what we did in here relate to the conditions in the backyard?
2. Aden: It relates because there umm ... the conditions, well it-it doesn't relate because

the conditions in here... say-let's say it may grow that tall right now in that thing (*looking at light boxes in the room*) right in that box thing but out in the backyard it, it will be way taller.

3. Mrs W: You think it will grow more in the backyard than with the Wisconsin Fast Plant™.
4. Aden: Cause that box doesn't give that much light but the sun, it gives a lot of light.

In this episode, Mrs. W. asked students to remind her about the mappings between the backyard and experimental conditions. Analysis suggested that she treated these mappings as unproblematic and was seeking to review connections (e.g., both had sunny conditions). However, Aden brought up a difference between the two systems, stating that the light was stronger in the backyard than in the experiment, and indicated a result of the difference, that the same plant grown in the backyard would be taller than the specimens in the classroom experiment. These kinds of conversations were scattered throughout students' work with the experiment, suggesting that students noted the slippages between the experiment and backyard and thought they had consequences for comparing results in the two systems.

Near the end of the investigation, the use of mappings exploded (ten of the 16 episodes occurred on two consecutive days of instruction; eight of these involved student initiation of relations or implications). Mrs. W. introduced a claim about the backyard based on the results of the experiment. She argued that the class had shown that the best amount of light for plants was sun, and that therefore

> "I think the just right amount of light for all plants in the backyard is sun. So when we go outside, I think we will find no plants in the shade, some plants in the sun and shade, and lots and lots of plants in the areas that always get sun."

As she presented her argument, several students began to disagree with her. Initially, students noted that her argument was not correct based on what they had seen in the backyard, calling out "No, mine's not really in a place in the sun" and "Because when you go in the Wild Backyard there are some- there are some plants…in the wild backyard, but in HERE they're not growing" Students then began to generate reasons for the differences in plant growth in the two locations. For example, Steven argued, "the lightbox doesn't have as much sun as the sun, we're just pretending it does" while Madison suggested that the shade outdoors was "not always in the shade because sometimes it is in the sun (e.g., when the sun moves throughout the day)." Here, students explicitly recognized and responded to resistances, in that they argued that the results of the experiment did not mirror what they saw or would expect to see in the backyard setting it was meant to represent.

As they continued conversations in small groups the next day, several students noted that the experiment had used Wisconsin Fast Plants™, while there were many kinds of plants outdoors, as when Azhad argued "We have two different plants…some are MADE to live in the shade," prompting Mrs. W. to revoice his contribution, "OK, so you're saying these plants aren't like all plants," initiating the following conversation.

1. Ellen: No they're not because they might [come from different countries. Different cities. Different kinds of *undecipherable*]
2. Azhad: [OK. Do you have any] plants next to your bush?
3. Mrs. W: Next to what bush?
4. Azhad: Your bush.
5. Mrs. W: At home?
6. Azhad: Yes.
7. Mrs. W: Yeah, I do.
8. Azhad: And I have plants growing under MY bush. But- cause they're different plants. I have roses, I've got daisies.
9. Mrs. W: But what makes the difference?
10. Azhad: (*Points to lightbox*.) Cause these are different plants. These are Wisconsin Fast Plants. [Mine is daisies]
11. Mrs. W: [So you mean] different plants need different amounts of light?
12. Ellen: Yes [cause] they don't [really] need the same thing
13. Azhad: [Yes]
14. Jasmine: [Like] wild strawberries they hardly need any light because they're growing right there in the shade.

In this excerpt, all three students talking with Mrs. W. suggested that the Wisconsin Fast Plants used in the experiment could not stand in for all plants. Both Azhad and Jasmine introduced examples of plants growing in

the shade (Line 8 and Line 14). In doing so, they grappled with resistance that they experienced: plants in other settings growing in conditions that their experimental plants could not. In response to these resistances, they developed accommodations by using mappings to support nascent model-fit practices, suggesting that differences in plant kind could account for differences in growth.

Seeing the results in both systems (experimental and backyard) appeared to establish a context in which considered how the experiment was and was not a useful model of the backyard setting. When students' attention was directed to the question of whether the experiment could predict growth in the backyard, they engaged more fully and heatedly in discussing mappings between the two systems, initiating new relations (e.g., plant kind, moisture) and participating in longer episodes with more widespread participation. Therefore, here again, resistance supported students to develop a practice that was useful and meaningful to them.

Mapping as a Site for Conceptual Work

Engagement in mapping between the two systems demonstrated similar forms of conceptual opportunities, in that students *differentiated ideas* and *called on ideas as mechanisms.* For example, to argue against Mrs. W., students differentiated plants into "kinds of plants" as they argued that Wisconsin Fast Plants were not like all plants and introduced "daisies" (Line 10) and "strawberries" (Line 14). They differentiated growth conditions as they began to focus on how much light the plants received and whether moisture differences might also matter for the distribution of plants outside. They also evoked mechanisms to justify the relevance of the differences they noted. As students suggested that it was important that the Fast Plants were a different kind of plant than those in the backyard, they began to talk about plants' needs, supporting their claim that plant kind mattered by using ideas of differential success, evident in Azhad's statement that "some are MADE to live in the shade" and Jasmine's explanation that "wild strawberries they hardly need any light" (Line 14). The identification of needs allowed Mrs. W. to guide students toward thinking about why different plants might have different needs, provoking talk about plant structures and strategies and introducing a book that provided new information. When Diego argued that the differences in results were caused because the backyard got more water than the Fast Plant systems, students questioned why he thought water could make up for lack of sunlight.

Here again, resistances situated the interrelated development of a practice, mapping between the two systems, and concepts, namely differentiation of conditions and explanations of differential success. Ideas such as plant kind, amount of light, or presence of moisture were not useful to students when they were discussing the experiment in the absence of considering the backyard. However, differing results across the two systems destabilized ideas that had been effectively black-boxed by the experimental conditions, causing students to develop forms of accommodation consisting of both mapping practices and new categories and explanations.

Discussion and Conclusions

In this paper, I described two activities, *defining attributes* and *mapping between the experiment and target system,* in which students engaged as they conducted the plant growth experiment. The results suggest that these activities were constituted as scientific practices in this classroom community. They appeared to be meaningful to students, in that they were often initiated by students rather than teachers and they served identifiable purposes in their work: coping with seeing the same thing when plants were changing and understanding why the results of the plant growth experiment did not represent growth patterns in the backyard. Described more generally, these functions, seeing the same thing as others and mapping between experiments and phenomena to evaluate model-fit, are central to scientific activity (Gooding, 1990; Nersessian, 2012; Pickering, 1995).

The results support the conjecture that students' scientific practices would emerge and be refined in response to resistances in the material system. For example, students engaged in definition in order to agree on plant features in the face of change, a resistance that made it difficult to agree on what attributes were and what they meant. Their use of the practice was related, but not identical, to that of teachers, who initiated episodes of defining to help students refine ideas that, from the teachers' point of view, appeared vague (e.g., "big."). An additional finding is that classroom structures and actions were important design features that made these resistances visible and problematic. Students were repeatedly asked to present claims about plant success and note attributes that supported their ideas, making variability in their interpretations visible and seeding definition. The teacher purposefully introduced a problematic claim (that there should be no plants in the shady areas outside) in order to highlight a resistance; this action supported an explosion of mapping practice.

In addition, these results suggest that purposefully designing resistances into students' work can support the integration of scientific ideas and scientific practices in instruction. First, the paper highlights the conceptual affordances of the very parts of experimental activity that are usually simplified for use with young students. For example, many studies have shown that students do not "see" what scientists see when looking at phenomena (Chinn & Malhotra, 2002; Eberbach & Crowley, 2009). As a result, young students are often presented with categorical variables or provided explanations that essentially tell them what to see. Here, however, wrestling with what to see and how to see it in the same way was both an accessible activity for students and a site for conceptually rich talk about plant life cycles, an idea students found challenging in the

backyard setting. Likewise, dealing with differences in the two systems both seeded model-fit practices and provided an opportunity to further differentiate ideas about light and plant kind. The paper also contributes to the literature by describing three forms of "conceptual development" that occurred as students developed new practices to cope with resistances: differentiating categories, relating entities or attributes, and calling on mechanisms. Future research will focus on predicting the conceptual affordances of particular resistances and preparing teachers to recognize the emerging opportunities for students to differentiate, relate, and call on ideas as mechanisms. In this way, resistance can be made into an affordance, rather than a source of chaos.

Finally, and in keeping with the theme of the conference, this paper highlights a distinction between engagement in "practice" and "practices" to which the field might profitably attend. Here, practices were lent meaning by students' engagement in *scientific practice,* in that they were actively wrestling with developing shared ways of seeing and knowing in the face of resistances. One fruitful direction for future work might be to make *practice* a central target of design, with the understanding that epistemic *practices* are meaningful only in the context of epistemic struggles.

Endnotes

(1) "We" is used to refer to the author, the larger research team, and the classroom teacher.

(2) "Teachers'" refers to the classroom teacher and the author. Since both of us asked students questions and commented on their ideas, I treated both of our comments as framing and elaborating activity in ways consistent with "teaching."

(3) Transcript conventions: CAPS emphasis; [ ] overlap; - self interruption; … pause; (*italics*) gesture; other punctuation added to increase readability.

References

Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education, 95*(2), 191-216. doi: 10.1002/sce.20420

Chinn, C., & Malhotra, B. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded? *Journal of Educational Psychology, 94*(2), 327-343.

Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Sciences, 10*(1/2), 113-163.

Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research, 79*(1), 39-68.

Ford, M. J. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education, 92*(3), 404-423.

Gooding, D. (1990). *Experiment and the making of meaning*. Dordrecht: Kluwer Academic Publishers.

Gresalfi, M., Martin, T., Hand, V., & Greeno, J. (2009). Constructing competence: An analysis of student participation in the activity systems of mathematics classrooms. *Educational Studies in Mathematics, 70*(1), 49-70.

Hall, R., & Greeno, J. G. (2008). Learning and understanding concepts in practice. In T. Good (Ed.), *21st century education: A reference handbook.* . Thousand Oaks, CA: Sage.

Hogan, K., & Corey, C. (2001). Viewing classrooms as cultural contexts for fostering scientific literacy. *Anthropology & Education Quarterly*, 214-243.

Kuhn, D., & Pease, M. (2008). What needs to develop in the development of inquiry skills? *Cognition and Instruction, 26*(4), 512 - 559.

Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development, 23*(4), 512-529.

Manz, E. (2012). Understanding the codevelopment of modeling practice and ecological knowledge. *Science Education, 96*(6), 1071-1105. doi: 10.1002/sce.21030

National Research Council (2012). *A framework for K-12 science standards: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: The National Academy of the Sciences.

Nersessian, N. J. (2012). Engineering Concepts: The Interplay between Concept Formation and Modeling Practices in Bioengineering Sciences. *Mind, Culture, and Activity, 19*(3), 222-239.

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. Chicago: University of Chicago Press.

Roth, W., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30*(2), 127-152.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Acknowledgements

This work was supported by an IES Predoctoral Fellowship and National Science Foundation Grant 0628253.

An Analytic Tool for Supporting Teachers' Reflection on Classroom Talk

Gaowei Chen, University of Hong Kong, Pokfulam Road, Hong Kong, gwchen@hku.hk
Sherice N. Clarke and Lauren B. Resnick, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA
Email: sclarke@pitt.edu, resnick@pitt.edu

Abstract: Teachers can reflect on and analyze their classroom talk to inform their instructional practice. When teachers try to do so however, they often face analytic difficulties regarding the data set (data input, data transformation, and utterances by unknown speakers), coding (coding complexity, reliability, and efficiency), visualization (representations of a variety of information, synchronization of displays, and adaptation to the data/codes changes), and tracking and comparison (many students' actions across discussion sessions). This paper introduces an analytic tool called classroom discourse analyzer (CDA) to address these difficulties, as shown in the analyses of classroom discourse from a fourth grade science class. The analyses demonstrate how CDA can be used by teachers to support their reflection on classroom talk and how it can provide personalized, data-supported evidence to inform teachers' classroom practice.

Introduction

Teachers play a crucial role in classroom discussions. Their orchestration of student participations, argumentation and evaluations can affect students' learning processes and outcomes (Resnick, Asterhan, & Clarke, in press). However, currently teachers' performance in helping student learn through discussion is far from satisfactory (Mercer, Dawes, & Staarman, 2009; McNeil & Pimentel, 2010). Given the large benefits of effective classroom talk on students' learning, teachers often find it difficult to engage students into productive discussions involving deep reasoning and argumentation (Clarke et al., 2013; Howe & Abedin, 2013; Pimentel & McNeil, 2013). This raises the challenge of teacher education and professional development (PD) on their classroom discussions with students.

To address this issue, many PD programs have developed instructions to guide teachers' reflection on their classroom interaction. Although reflection has been widely accepted as a useful means of learning in PD programs (Korthagen, & Vasalos, 2005; Pollard et al., 2008), the format of reflection from one's memory has been questioned by some researchers who argued that one's memory cannot always be reliable, neither can it notice and memorize every detail of classroom interaction (McCullagh, 2012; Rosaen, Lundebrg, Cooper, Fritzen & Terpstra, 2008). As such, more recently a growing number of researchers and practitioners have been investigating the use of videos as a facilitator in teacher education and PD (e.g., Baecher, Kung, Jewkes, & Rosalia, 2013; Borko, Koellner, Jacobs, & Seago, 2011; van Es, 2012). Studies have compared various types of video and got mixed research results. For instance, analyzing videos of one's own teaching was shown to be more effective regarding activation experience but less effective regarding emotional/motivational involvement than analyzing others' video (e.g., Kleinknecht & Schneider, 2013). The mixed results encouraged us to explore alternative means of PD that can not only take the advantage of using authentic classroom data as video reflections do, but also provide teachers with data-supported evidence to inform their practice.

The emerging fields of learning analytics provide ample opportunities for the employment of data as evidence for teaching and learning (1st International Conference on Learning Analytics & Knowledge; Baker & Yacef, 2009; Siemens et al., 2011). In particular, discourse analytics in which language is a primary tool for knowledge negotiation and construction (Shum & Ferguson, 2012) allows teachers to identify classroom discussion behaviors and patterns through the analysis of discourse data. As data and analytics are reshaping the way of teaching and learning, tools that can interact with teachers to visualize and track classroom discourse and communicate analytic results should be explored to meet the needs of teachers.

This paper introduces a novel discourse analytic tool called Classroom Discourse Analyzer (CDA) for teachers. We explicate how CDA addresses the difficulties regarding data set, coding, visualization, and tracking and comparison that teachers face in analyzing their classroom discourse data. Data set difficulties include data input, data transformation, and dealing with utterances by unknown speakers. Difficulties involving data coding include the coding of complex classroom discourse, coding reliability, and coding efficiency. The visualization issues include the visual representations of many things, synchronization of different displays, and adaptation to data/codes changes. Lastly, there are also difficulties regarding tracking and comparing teacher's and students' actions across multiple sessions.

We showcase CDA by applying it to the discourse data from a fourth grade science class. The data was first video recorded and transcribed by humans. The female teacher and 16 students (9 males, 7 females)

 © ISLS

contributed 1,939 teacher turns and 1,926 student turns in 30 discussion sessions (total duration: 8 hours and 51 minutes). The analyses and visual representations show how CDA can support teachers' reflection on their classroom discussions and how it can provide personalized, data-supported evidence to inform teachers' classroom practice.

Classroom Discourse Analyzer

When trying to reflect on and analyze classroom discourse, teachers often face analytic difficulties regarding the discourse data set, coding, visualization, and tracking and comparison. Table 1 presents the difficulties as well as what strategies that CDA takes to address these difficulties. We classify the difficulties into four categories (i.e., data set, coding, visualization, and tracking and comparison) and discuss below how CDA addresses each of them.

Table 1. Difficulties for teachers in analyzing classroom discourse and CDA strategies.

Analytic Difficulty	**Classroom Discourse Analyzer Strategy**
Data set	
• Data input	• Simple input template (session, turn, speaker, content)
• Data transformation	• Automated transformation into data of three levels (turn [teacher turn vs. student turn], exchange [teacher-students turn-taking], session [sequences of exchanges])
• Unknown speakers	• Naming array (*individual student* [e.g., *John*; identifiable student], Male [a male student], Female [a female student], SS [a group of students], or S [an unidentifiable student])
Coding	
• Complexity of classroom discourse	• Multi-dimensional coding with mutually exhaustive, exclusive categories
• Coding reliability	• Multi-dimensional coding with mutually exhaustive, exclusive categories
• Coding efficiency	• Automated coding by machine learning software
Visualization	
• Displays of many participants, turns and codes	• Visual displays of many students, many turns, and multiple codes by different colors/shapes and in different frames
• Synchronization of displays	• Automated synchronization of various displays
• Adapting to data/codes changes	• Real-time link of visualization to data and codes
Tracking and comparison	
• Many students' actions over time	• View of anyone's (teacher or students) any actions (e.g., number of words, new ideas, questions) in any discussion sessions

Data Set

CDA addresses the discourse data set issues (data input, data transformation, unknown speakers) with a simple input template, automated transformation, and a student naming array. The input of classroom discourse can include various entries and may take various forms. To minimize the workload of teachers, a simple input template uses four variables to represent a conversational turn, which are session (i.e., session # in a series), turn (i.e., turn # in a session), speaker (i.e., name of the speaker), and content (i.e., entire content of a turn). The input can be any number of sessions, turns, or speakers.

Classroom discourse data often need to be transformed into right types of data sets for subsequent analyses. CDA can automate the data transformation processes, which transforms classroom discourse into three-level data with a nested structure, namely turn, exchange, and session levels. Turns are nested within exchanges and exchanges are nested within sessions. At the turn level, all conversational turns are classified into teacher turns or student turns. At the exchange level, the sequences of teacher student turn-taking are identified (e.g., a sequence of "T->S_1->T->S_2" or "T->S_1->S_2->T"). Lastly, at the session level, the boundaries between any two consecutive discussion sessions are automatically detected.

Classroom discussions often involve a large number of participants (teacher and many students). This may create obstacles for the identification of the participating students in some utterances, especially with audio-taped data. To address this issue, CDA provides a student naming array to classify the speakers into one of the five following categories: *individual student* (e.g., *John*; identifiable student), Male (a male student),

Female (a female student), SS (a group of students), or S (an unidentifiable student). According to the naming array, student speakers are labeled as "S" only if they cannot be classified into the any of the first four categories. This strategy is to retain the relevant information about the speakers for subsequent analyses.

Coding

CDA addresses the coding issues (i.e., complexity of classroom discourse, coding reliability, coding efficiency) with multi-dimensional coding at the unit of conversational turn and the use of machine coding. The complexity nature of classroom discourse often requires a coding framework with many categories. As the number and complexity of categories rise, the training time for teachers and the overall coding time rise, coding conflicts rise, and coding reliability and efficiency fall (Chiu & Khoo, 2005).

By using multi-dimensional coding at the unit of conversational turn, CDA can reduce the number of needed variables, increase coding reliability, and capture the discourse data's complexity. For example, CDA provides a three-dimensional framework to code a student turn. The three dimensions are: evaluation (with "agree", "disagree", and "neutral" as the categories), knowledge content (with "new idea", "repetition", and "no academic content" as the categories), and invitation to participate (with "statement", "question", and "command" as the categories). Because each dimension has three categories, this framework can capture 27 (3×3×3) different types of action. By coding one dimension at a time, a teacher/coder uses clear criteria to choose among only three possible codes, instead of 27. Such categories are mutually exclusive, exhaustive and sufficiently comprehensive to characterize classroom discourse. Meanwhile, the simplification can reduce the coding complexity and likely increase inter-coder reliability.

Coding classroom discourse is often laborious, especially for large data sets. It often needs at least two coders for checking the inter-coder reliability (Krippendorff, 2004). CDA supports the use of automated coding for coding efficiency. Computer coding can be based on either a set of fixed decision rules or human codes for similar data (Erkens & Janssen, 2008; Rosé et al., 2008). For example, lightSIDE can train machine coding models based on a sample of human codes and meanwhile provide a series of coefficients (e.g., reliability test) for measuring the models' performance. Satisfactory models can then be used to code the discourse data automatically and the coding results can be entered into CDA for subsequent analyses.

Visualization

Visualization with the synchronized discourse transcripts is an important function in CDA because it supports teacher reflection of classroom interaction by providing an activating experience similar to that using video as a facilitator for reflection (which is very often referred to as a 'vivid secondhand" experience; Miller & Zhou, 2007; Seidel et al., 2011). CDA can visually display participants, turns, and codes by a variety of shapes and colors, in different frames of the same visible window, and more importantly with a real time update to the changes of data/codes.

In more detail, first as classroom discourse often involve many participants (teacher and many students), many turns, and multiple codes (e.g., codes measuring turns, words, words per turn, teacher-student turn-taking, teacher turn attributes, and student turn attributes), the visual displays of them can be difficult. CDA addresses this issue by using different shapes/colors to represent these components and visualize them in different frames of the same window, so that they do not overlap with each other and can be viewed at the same place. For example, CDA uses the size of a bubble to represent the number of words in a turn, while the color of the bubble represents a particular code (e.g., a new idea) of the turn.

Second, teachers may get lost when navigating between different displays within CDA. To address this issue, CDA's displays in different frames are always synchronized. A teacher's activities in one frame (e.g., zooming in, zooming out, or clicking on a data point) will be automatically synchronized in all other frames. Third, a discourse data set and its codes might change from time to time, which requires visualization to be automatically and dynamically linked to the changes. In CDA, any changes to the participants, turns, or codes will be updated in the visualization instantly. Therefore, no matter what changes teachers have made into the existing discourse data set, the visualization of participants, transcripts, and codes are updated automatically.

Tracking and Comparison

To obtain a deeper understanding of the classroom discourse, teachers often need to interact with the visualization to track individual students' talk over time or compare teacher-student interactions in different sessions. Due to the complex nature of classroom discourse, the visualization often includes many students' many actions (e.g., number of turns, number of words per turn, evaluation of previous turn [agree, disagree, or neutral], knowledge content [new idea, repetition, or no academic content], invitational form [statement, question, or command]) across sessions. Such visual information can sometimes be overwhelming so it might hinder teachers from finding interaction changes across discussion sessions.

To facilitate the tracking and comparison, CDA allows a teacher to select any participants' (teacher and/or students) any actions in any sessions. For example, a teacher may choose to only display a particular

student's turns according to time sequence, in order to see if the student is participating in the discussions more or less frequently over time. Alternatively, a teacher may select to view all students' talk in two different sessions to observe if these students participated in the discussion of the two topics differently. Once a trend or difference is identified, teachers can zoom in to view the visualization in further details.

Showcasing Classroom Discourse Analyzer

In this section, we showcase CDA by applying it to the transcripts from a fourth-grade science class. After describing the data set and coding, we show in what ways that teachers may use CDA to analyze the discourse and thereby inform their classroom talk with students.

Data

The data set includes 1,939 teacher turns and 1,926 student turns transcribed from videotapes of 30 discussion sessions in a 4th grade science class. The participants were a female teacher and 16 students (9 males, 7 females). The total time of discussion was about 8 hours and 51 minutes. Sample discussion topics include: "what causes the water level to rise?"; "Same volume, same weight?"; "How can we measure the volume of a liquid?". The format of the input data in CDA is presented in Table 2. Teachers only need to provide data along the following four columns: session, turn, speaker, and content.

Table 2. Format of the input data in CDA.

Session	Turn	Speaker	Content
1	1	T[a]	What is an earth material? Amalia, what's an earth material?
1	2	Amalia[b]	Materia- Material is like, um, like well, you could say like what's under our feet, you could say like um, maybe like soil and rocks.
1	3	T	OK. What's another way you would describe earth materials? What does that word or that term mean to you? Louie?
.	.	.	…
30	87	T	And what do we call that?
30	88	SS[c]	Water displacement.
30	89	T	I could use my water displacement again to get the exact measurement, because that worked for us with the rock, so I believe that would probably work for us with the gravel.

[a] Teacher. [b] Student name was pseudonym. [c] A group or whole-class of students.

Coding

CDA used multi-dimensional frameworks for coding the teacher and student turns. (Teachers may also create their own coding framework in CDA). CDA automatically computed the number of words in a teacher or student turn and identify teacher student turn-taking patterns over time (e.g., T->S_1->T->S_2->T vs. T->S_1->S_2->T).

The teacher turns were coded based on the scheme of teacher Accountable Talk® during the discussion, which includes eight categories of Accountable Talk moves (i.e., "say more", "revoice", "press for reasoning", "challenge", "restate", "add on", "agree/disagree", and "explain other"; Resnick, Michaels, & O'Connor, 2010). For example, to encourage individual students' thinking, teachers can ask students to say more (e.g., "Can you say more about that?") or to explain their reasoning (e.g., "Why do you think that?"). To facilitate students' thinking with others, teachers can ask students to extend the ideas of others (e.g., "Who can add onto to that idea?") or to evaluate someone else's reasoning (e.g., "Do you agree / disagree?"). Machine coding software (lightSIDE; Mayfield & Rosé, 2013) was used to facilitate the coding of teacher turns.

The student turns were coded using a multi-dimensional framework: a) knowledge content (new idea [with justification or not], repetition, or no academic content), b) evaluation of previous turns (agree, disagree, or being neutral), and c) invitation to participate (statement, question, or command; Chen & Chiu, 2008; Chen, Chiu, & Wang, 2012).

Analyses

We first introduce the overall interface of CDA. We then show how the various displays in CDA help visualize the discourse processes and how they can inform teachers about their talk with students in the classroom.

CDA Interface

Figure 1 shows an overview of the classroom discourse from the 30 sessions. The user interface of CDA includes 5 frames for visualization (see Figure 1). Frame 1 is the place that visualizes the distributions of the

teacher and student turns. The vertical axis is participants and the horizontal axis is timeline. Each line of bubbles represents a student's or the teacher's (the dark red bubbles) talk. Size of bubbles represents the number of words in a turn. The teacher's line of bubbles near the bottom of the frame 1 shows that she talked much more than the students across the 30 sessions.

Frame 2 represents teacher-student turn-taking patterns automatically generated by CDA. The vertical axis is the number of students between two adjacent teacher turns and the horizontal axis is timeline. Frame 3 shows summary statistics for each row of data (e.g., teacher's and student's percentages of total words in the 30 sessions). The transcripts are displayed in frame 4 and the coding frameworks for teacher and student turns are displayed in frame 5. The visual displays in different frames are always synchronized to facilitate viewing. For instance, Figure 1 shows that a particular turn (turn 38 in session 16 by the student Johnny, a pseudo name in this study) can be easily located in frames 1, 2, and 4 simultaneously.

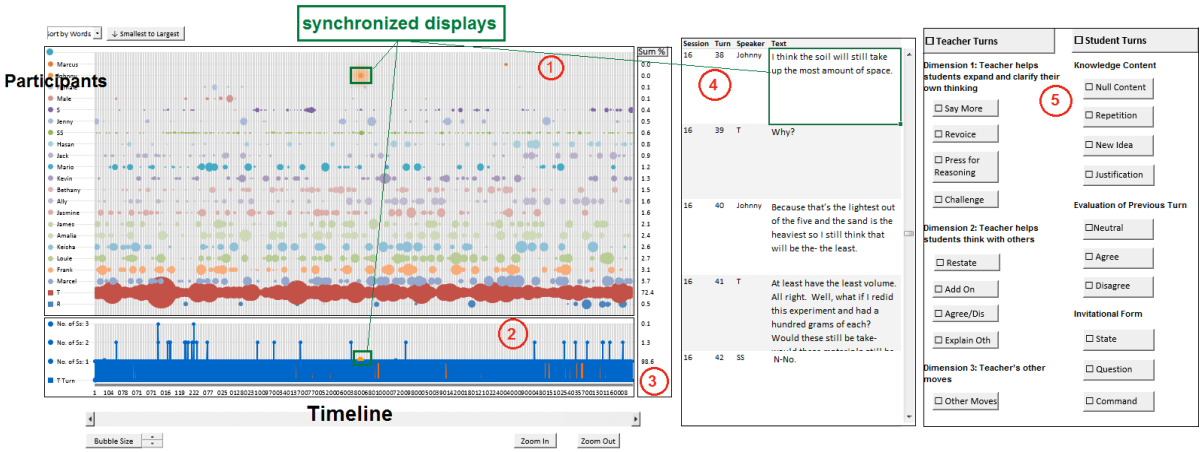


Figure 1. An overall interface of CDA.

Distributions of Teacher and Student Turns in a Session

In addition to the overview of multiple sessions, teachers can use CDA to visualize the speakers' turns and words in a particular session. Figure 2 shows the teacher's and 14 students' turns and words in session 26. (Two students were absent from the session). The discussion topic is "what causes the water level to rise?" There were 71 teacher turns and 70 student turns in the session. The session lasted about 20 minutes and 47 seconds. As shown in Figure 2, the students were sorted by the amount of words this time. While *Mario* only spoke one word in one turn, *Marcel* spoke 233 words in two turns in the session. Based on the information the teacher may balance students' participations in subsequent discussions (e.g., inviting *Mario* to participate more in future sessions).

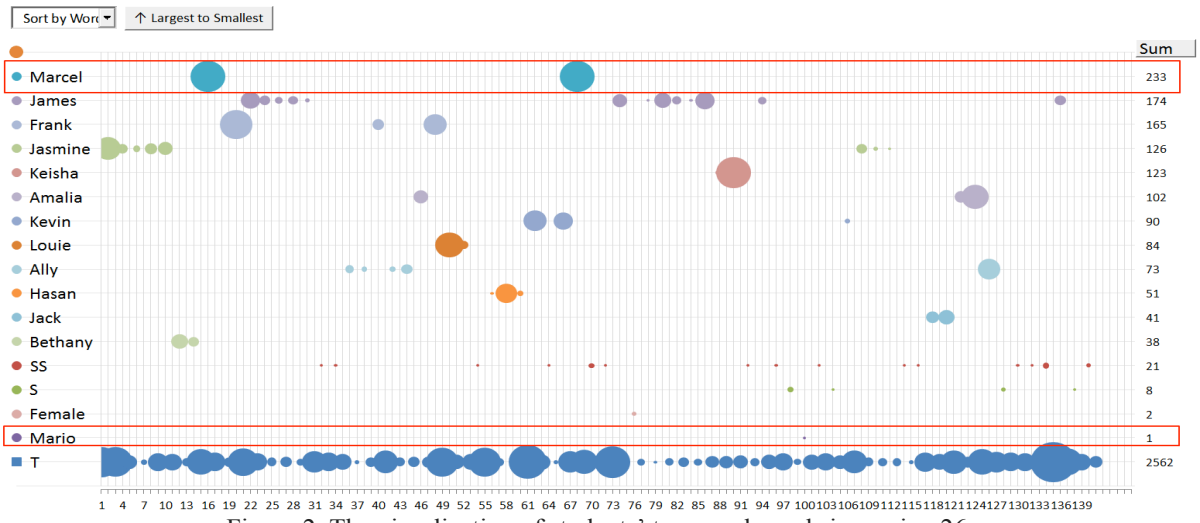


Figure 2. The visualization of students' turns and words in session 26.

Teacher-Student Turn-Taking Patterns in a Session

CDA can automatically generate the visualization of teacher and students' turn-taking patterns during a discussion. There are three major patterns during classroom discussions. The first is teacher's talk with a

particular student for two or more turns (or "T->S_1->T->S_1"). Another is teacher's talk with one student for one turn and then turn to another student (or "T->S_1->T->S_2"). The last one is that teacher allows two or more students to talk with each other before taking the turn (or "T->S_1->S_2->T").

See Figure 3 for the turn-taking patterns in session 26. During this discussion, the teacher and students used many "T->S_1->T->S_1" patterns (dashed orange lines) and many "T->S_1->T->S_2" patterns (solid blue areas), but no "T->S_1->S_2->T" patterns. For example, Figure 3 shows a place where the teacher talked with a particular student for 11 consecutive turns before going to another student ("T->S_1->T->S_1->T->S_1->T->S_1->T->S_1->T"). This might be the place that the teacher helped the student clarify or expand his or her own thinking. At another place, the teacher went through eight students and talked with each student with only one turn ("T->S_1->T->S_2->T->S_3->T->S_4->T->S_5->T->S_6->T->S_7->T->S_8->T"). This might be the place where the teacher helped students think with one another. Teachers can click on the visual representation to view the corresponding video/transcripts details and purposefully examine whether a teacher has managed the discourse segments appropriately or whether he/she missed opportunities for fostering student thinking in that context.

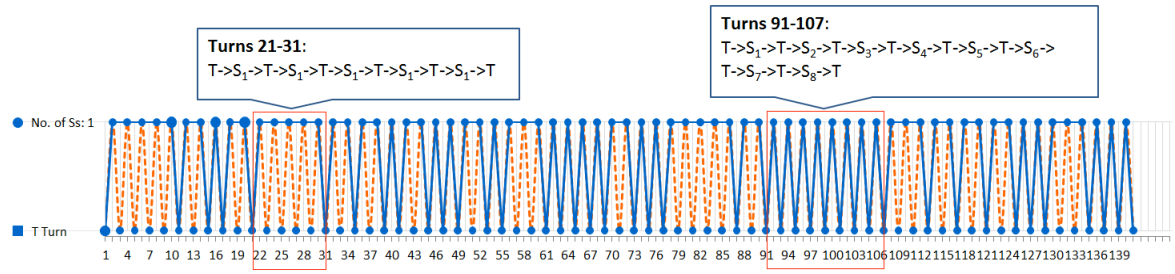


Figure 3. The visualization of teacher-student turn-taking patterns in session 26.

Tracking Teacher and Student Actions in a Session

Teachers can use CDA to track teacher's or any individual students' particular actions during classroom discussions. Figure 4 shows the visualization of a) the teacher's use of "agree/disagree" moves, and b) the students' use of "disagreement" moves in session 26. The purpose was to track how students disagreed with one another during the discussion. From Figure 4 we can see whether they disagreed with one another as a result of the teacher's probes or they disagreed with one another voluntarily.

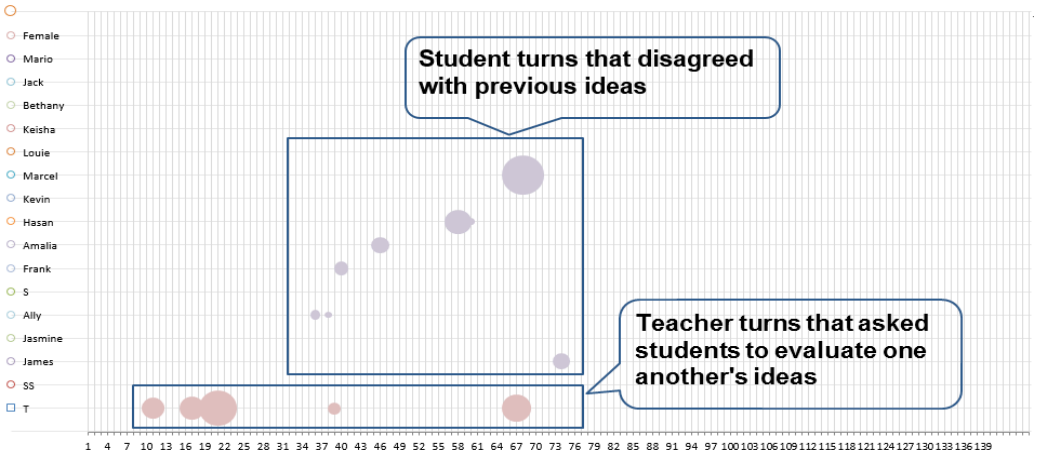


Figure 4. Visualization of teacher's "agree/disagree" turns and students' "disagreement" turns in session 26.

As shown in Figure 4, there were 5 teacher turns that asked students to evaluate one another's ideas and 8 student turns that disagreed with one another's ideas. What is interesting here is that only 2 of the 8 disagreement turns were elicited by the teacher. The other 6 turns were not preceded by teacher "agree/disagree" turns, suggesting that these students were likely to disagree with the ideas by their peers explicitly during the discussion (McNeill & Pimentel, 2010). Teachers may zoom in to reflect on why the students were having conflicting views and how the teacher moderated the disagreements in this episode.

Comparisons Between Classes

Moving beyond a single class, teachers can use CDA to compare their classroom discourse data with those from other teachers. This is a useful function especially in teacher PD because it can help teachers learn from the analyses of one another's classroom talk. Figure 6 shows the teacher-student turn-taking patterns of a discussion

segment from a ninth grade biology class (discussion topic: Punnett square; one male teacher and 16 students; duration: 16' 00"). This visualization is in sharp contrast to that in Figure 3. The ninth grade biology class used much fewer "T->S_1->T->S_1" patterns (dashed orange lines), but a lot more "T->S_1->S_2->T" patterns (higher, blue lines with a peak) than the fourth grade science class. As can be seen in Figure 6, there were as many as five students taking the turns to talk between two teacher turns (e.g., "T->S_1->S_2->S_3->S_4->S_3->S_5->T" as shown in the rectangles). The visualization showed that students in this discussion session had a lot of opportunities to interact with one another directly, which indicates a presence of students' dialogic interactions in the session (Lehesvuori et al., 2013; McNeill & Pimentel, 2010; Schultz & Oyler, 2006).

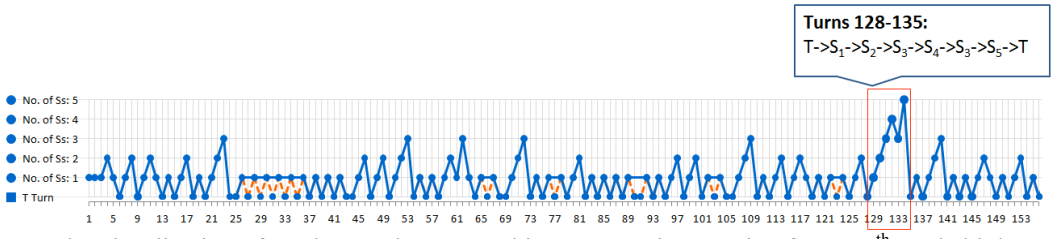


Figure 6. The visualization of teacher-student turn-taking patterns in a session from a 9th grade biology class.

Discussion and Conclusions

In this paper, we explicated how teachers can use CDA to analyze their own classroom discourse. While many difficulties encumber teachers' analysis of their discourse data, CDA addresses each of them. CDA can be used by teachers and teacher educators to support teachers' self-assessment and reflection of their classroom discussion processes and provide personalized, data-supported evidence to inform their classroom practice. Through visualizing, tracking, and comparing classroom discourse processes, CDA can be applied to address questions regarding a) how classroom talk develop over time, b) how teacher and students take turns to interact with one another, and c) how their behaviors differ across sessions and classes.

The CDA results of the above case showed that it can be used to visualize teacher and student talk over time. This allows teachers to track any participants' actions (e.g., students' disagreements with one another, new ideas, justifications) in a discussion session. The information can also be used to inform teacher's guidance and feedback in future discussions. CDA can also automatically visualize teacher and students' turn-taking patterns to reveal at a glance the social structure of a discussion. For example, the occurrences of "T->S_1->S_2->T" pattern (two or more students talk between two adjacent teacher turns) might indicate the places where students voluntarily evaluate someone else's idea, raise a question, or propose a new idea (McNeill & Pimentel, 2010; Schultz & Oyler, 2006). Moving beyond a single session, teachers can use CDA to compare students' behaviors in different sessions. Furthermore, a group of teachers can use CDA to view and compare the discourse from one another's classes, which would also help develop a learning community among teachers.

There are challenges that might limit the application of CDA in practice, which include data collection, transcribing, and coding. All the three processes, especially data transcribing and coding, are laborious and expensive, thereby hindering the rapid feedback and continuous support that CDA can provide. The data of classroom discussions are often collected by audio or video taping (e.g., the data used in the above case). In a classroom discussion that involves a large number of students, it is sometimes difficult to capture the information of who speaks at what time. Alternative ways of recording classroom discussions, such as microphone-array systems (Sun & Canny, 2012), may help address this issue.

The data in the above case were all transcribed manually from videotaped records. The process was often time and labor consuming. New technology in the areas of speaker identification and speech recognition may be used to facilitate the transcribing of classroom discussions (e.g., Vandyke, Wagner, & Goecke, 2013; Walker et al., 2004). Whether computer transcribing is comparable to external human transcribing remains an open research area. Another strategy is that teachers can select to transcribe the small segments that they are interested. (Sampling portions of the data omits substantial data, which can bias the analysis results.)

Coding transcripts of classroom discussions also takes time and effort. It has been showed that computer coding was reasonably acceptable as it coded for some particular categories (Clarke et al., 2013; Mayfield, Laws, Wilson, & Rosé, 2013). In the above case, we have used the machine learning software lightSIDE (Mayfield & Rosé, 2013) to help code for some teacher talk moves that do not require the teacher to be able to paraphrase or interpret the student speaker's "meaning." They are the moves that mainly use general expressions (e.g., press for reasoning: "why do you think that?"; add on: "who can add on?"; agree/disagree: "do you agree/disagree?"). Whether computer coding can be used to code complex categories of classroom talk remains an open research area.

References

Baecher, L., Kung, S. C., Jewkes, A. M., & Rosalia, C. (2013). The role of video for self-evaluation in early field experiences. *Teaching and Teacher Education, 36*, 189-197.

Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009. *Journal of Educational Data Mining, 1*(1), 3-17.

Borko, H., Koellner, K., Jacobs, J., & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM, 43*(1), 175-187.

Chen, G., & Chiu, M. M. (2008). Online discussion processes. *Computers & Education, 50*(3), 678-692.

Chen, G., Chiu, M. M., & Wang, Z. (2012). Social metacognition and the creation of correct, new ideas. *Computers in Human Behavior, 28*(3), 868-880.

Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes: Dynamic multi-level analysis. *Small Group Research, 36*, 1-32.

Clarke, S. N., Chen, G., Stainton, C., Katz, S., Greeno, J. G., & Resnick, L. B. et al. (2013). *The impact of CSCL beyond the online environment*. Paper presented in the CSCL 2013 conference. Madison, WI.

Erkens, G. & Janssen, J. (2008). Automatic coding of dialogue acts in collaboration protocols. *International Journal of Computer-Supported Collaborative Learning, 3*(4), 447-470.

Howe, C. & Abedin, M. (2013). Classroom dialogue. *Cambridge Journal of Education, 43*(3), 325-356.

Kleinknecht, M., & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teachers and Teaching*, 33, 13-23.

Korthagen, F., & Vasalos, A. (2005). Levels in reflection. *Teachers and Teaching*, 11, 47-71.

Krippendorff, K. (2004). *Content Analysis*. Thousand Oaks: SAGE Publications.

Lehesvuori, S., Viiri, J., Rasku-Puttonen, H., Moate, J., & Helaakoski, J. (2013). Visualizing communication structures in science classrooms. *Journal of Research in Science Teaching*.

Mayfield, E., & Rosé, C. P. (2013). LightSIDE. *Handbook of Automated Essay Evaluation: Current Applications and New Directions*, 124.

Mayfield, E., Laws, M. B., Wilson, I. B., & Rosé, C. P. (2013). Automating annotation of information-giving for analysis of clinical conversation. *Journal of the American Medical Informatics Association*, amiajnl-2013.

McCullagh, J. F. (2012). How can video supported reflection enhance teachers' professional development? *Cultural Studies of Science Education*, 7, 137-152.

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms. *Science Education,* 94(2), 203-229.

Miller, K., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. *Video research in the learning sciences*, 321-334.

Pollard, A., Anderson, J., Maddock, M., Swaffield, S., Warin, J., & Warwick, P. (2008). *Reflective teaching: Evidence-informed professional practice* (3rd ed.). London, New York: Continuum.

Resnick, L. B., Asterhan, C. A., & Clarke, S. N. (Eds.) (in press). *Socializing Intelligence through Academic Talk and Dialogue*. Washington, DC: American Educational Reserach Association.

Resnick, L. B., Michaels, S., & O'Connor, C. (2010). How (well structured) talk builds the mind. In D. Preiss & R. Sternberg (Eds.), *Innovations in educational psychology* (pp. 163-194). New York, NY: Springer.

Rosaen, C. L., Lundeberg, M., Cooper, M., Fritzen, A., & Terpstra, M. (2008). Noticing noticing. *Journal of Teacher Education*, 59, 347-360.

Rosé, C. P., Wang, Y. C., & Arguello, J. et al. (2008). Analyzing collaborative learning processes automatically. *International Journal of Computer Supported Collaborative Learning, 3*(3), 237-271.

Schultz, B. D., & Oyler, C. (2006). We make this road as we walk together. *Curriculum Inquiry*, *36*(4), 423-451.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations. *Teaching and Teacher Education, 27*(2), 259-267.

Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society, 15*(3), 3-26.

Siemens, G. et al. (2011). Open Learning Analytics. *Society for Learning Analytics Research*.

van Es, E. A. (2012). Examining the development of a teacher learning community. *Teaching and Teacher Education, 28*(2), 182-192.

Vandyke, D., Wagner, M., & Goecke, R. (2013). Voice source waveforms for utterance level speaker identification using support vector machines. In *International Conference on Information Technology in Asia, CITA*.

Walker, W. et al. (2004). Sphinx-4: A flexible open source framework for speech recognition.

Acknowledgments

This work was supported in part by NSF grant SBE-0836012 to the Pittsburgh Science of Learning Center and by NSF grant DRK12-0918435 awarded to TERC.

The Discourse of Creative Problem Solving in Childhood Engineering Education

Elise Deitrick, Brian O'Connell and R. Benjamin Shapiro
Center for Engineering Education & Outreach, Tufts University, Medford, MA, USA
Elise.Deitrick@tufts.edu, Brian.O_Connell@tufts.edu, Ben@cs.tufts.edu

Abstract: Researchers and teachers are increasingly in agreement that classrooms should adopt more open-ended, ill-structured, creative problem solving pedagogies (Kapur, 2008). However, we lack sufficient understandings of how to assess the variegated *outputs* of learning activities that afford students considerable discretion over what they will produce, and of the mechanisms through which group work can produce those outcomes. In order to understand how collaborative problem solving discourse shapes the creativity of collaborative products (as measured by the novelty of those products), we analyzed collaborative problem solving talk and the resulting products designed for fictional character by 9 groups of middle-school aged youth. We found that engaged responses to peers' proposed design ideas are predictive of novel solutions.

Introduction

Creativity has been gaining attention in education as an important skill for students in a variety of disciplines. Researchers are beginning to recognize the need to study the role of creativity in learning, and how creativity is related to other important phenomena. Creativity has been directly implicated by prominent theories of giftedness in young students (Renzulli, 2005; Sternberg, 2005). Creativity has been connected with improved performance and retention in mathematics (Van Harpen & Presmeg, 2013; Yuan & Sriraman, 2011), found to be beneficial to science understanding and literacy (Develaki, 2010; Webb & Rule, 2012), and enhances retention in music (Peterson & Madsen, 2010). This excitement extends to policy circles as well: The International Society for Technology in Education lists it first among their student standards and other organizations place similar importance to this desired virtue (ISTE, 2012; Davies et al., 2013). Most recently the Next Generation Science Standards (NGSS) have included engineering design, among other reasons, because "engineering offers opportunities for 'innovation' and 'creativity' at the K-12 level" (NGSS, 2013).

Collaboration among students is another goal – long held by learning scientists – that is receiving broad attention. NGSS (2013) exemplifies this by expecting that students learn about working in a team and developing communication skills, stating, "these skills are likely to be acquired when students engage in projects based on the science and engineering practices and core content." Connections between these values, creativity and collaboration, have been examined in many various studies from collaborative creativity as a desired learning outcome (Sullivan, 2011) to the resources and obstacles found in teacher team creative collaboration (Kurtzberg & Amabile, 2010) to more in depth and wider examinations of the topic (Eteläpelto & Lahti, 2008).

Theory

Researchers have proposed many definitions for creativity (Glück, Ernst, & Unger, 2002; Taylor, 1988). So many, in fact, that many have found it ineffective to narrow the definitions to one that is universally accepted (Saunders & Gero, 2002). This lack of a communal definition made creativity a nebulous catch-all buzz word in research. Instead of focusing on the vast number of concepts creativity could arguably cover, we have chosen a clear *aspect* of creativity often encompassed in these definitions: novelty (Saunders & Gero, 2002; Shah, Vargas-hernandez, & Smith, 2003). Novelty is "a measure of how unusual or unexpected an idea is as compared to other ideas" (Shah et al., 2003).

Researchers have developed two broad categories of instruments for measuring novelty: comparison and selection. When using a selection scale, student work is related to a set scale, as in a rubric. When using a comparative scale, the differences between artifacts are examined and scaled in comparison to one another (Merrill, Charyton, & Jagacinski, 2008). While most grading is currently done on a type of selection scale, e.g., a rubric, a problem arises when trying to do the same with creativity. In a classroom, there is a culture of borrowing and picking up ideas from peers. This may lead to a classroom set of solutions that look remarkably similar and on a selection scale, would be scored similarly. However, on a comparative scale, solutions are scored based on their differences. This means that creativity is scored locally, allowing for students to be assessed on their ideas, regardless of the common features that may have come about from classroom influence.

Collaboration, like creativity, does not have an accepted definition in the research world. "The broadest (but unsatisfactory) definition of 'collaborative learning' is that it is a situation in which two or more people learn or attempt to learn something together" (Dillenbourg, 1999). As two people attempt to learn, they share

 © ISLS

ideas and knowledge that have the potential to be taken up by the group, often after critique or discussion (Soller & Lesgold, 2007). Due to the uncertain nature of collaborative learning, there is "a general concern is to develop ways to increase the probability that… types of interaction [that trigger learning mechanisms] occur" (Dillenbourg, 1999). Dillenbourg (1999) categorizes these catalytic activities into Setup of Initial Conditions, Over-Specifying Collaboration Contract with a Scenario Based on Roles, Scaffolding Productive Interactions by Encompassing Interaction Rules in the Medium, or Monitoring and Regulating the Interactions.

There has been extensive research in the Computer Supported Collaborative Learning (CSCL) community around how to enable or study collaboration. This research examines what kinds of learning environment designs can support changes in the social organization of learning and enable youth to work together to construct new knowledge (Scardamalia, Bereiter, & Lamon, 1994; White, 2006). However, much of this work focuses on problem solving in domains where there are a priori knowable right answers. Learning scientists know relatively little about how collaborative discourse shapes solutions to open ended problems, particularly about how student talk can support the development of creative solutions to those problems.

Collaboration and creativity are two very complicated subjects, made up of a series of observable and unobservable factors whose relations the aforementioned studies have examined to varying levels of detail. What sort of discourse is associated with creative problem solving? Existing work in collaboration has shown engaged responses to peers' ideas can lead to correct answers in group problem solving, but this work has been limited to studying problem solving collaboration around problems that only have a single correct or incorrect solution (Brigid Barron, 2003). In this paper, we build upon Barron's (2003) methods to study problem solving discourse from a hands-on engineering summer camp to examine whether the characteristics of student discourse patterns that lead to success on closed-ended problems also predict more creative solutions in open-ended engineering design projects. Using mixed methods, we examine correlations between key features of students' collaborative discussion and the creativity of their solutions to open-ended engineering problems.

Study Context

We present data from a middle school age summer camp that was part of Integrating Engineering and Literacy at Tufts University's Center for Engineering Education and Outreach. This camp challenged participating youth to brainstorm, design, build, and test inventions that could assist fictional characters in children's literature who face a variety of problems. Three challenges, each based in a different book, were addressed over the span of three days. The camp had a morning session which used LEGO Mindstorms, and an afternoon session which used PaperBots, a newly developed educational robotics kit that was designed to be inexpensive and makes use of paper and craft material as building components (O'Connell, 2013). Each session of the camp consisted of 15 students in 4th through 6th grade, split into five groups of three.

Each challenge began with the participants and their teacher together reading a book. Then, the teacher-researcher asked participants to identify engineering problems within the story, facilitated by the lead instructor. Going back to their groups, participants then chose one of the identified problems and designed a solution to help the characters in the story using a robotics system. The data presented in this paper is taken from their interactions and solutions for the story Muncha Muncha Muncha by Candace Fleming, in which bunnies sneak into a farmer's garden and eat his vegetables at night despite his attempts to stop them. The children identified the problems of trying to help the farmer keep bunnies out of his garden or, alternatively, to help the bunnies get into the garden to eat the vegetables. After these two possible problems were identified, students returned to their groups, where they worked together to decide which problem to focus on. They then brainstormed possible solutions, and iteratively built and tested them. Students' group work lasted 2.5 hours, which was spread over two days. During this time, the teachers and researchers interacted with the groups, prompting them to talk about their ideas and what they were doing while working not to influence decision making. They did emphasize that their solutions had to work for the characters. The groups presented their solutions to the class at the end of the session.

Research Methods

We videotaped each group of students throughout their work, collecting about 115 hours of high definition video. Researchers deliberately avoided influencing solutions but intervened if significant group discord arose.

Computing Solution Novelty

We calculated novelty using a five step process: First, researchers identified *attributes*. Second, we assigned *weight*s to the attributes. Third, we mapped *ideas and features* to the identified attributes. Fourth, we computed *values* for ideas. Finally, we calculated a *novelty score* for each artifact.

Attributes were identified that were common between both populations and identifiable as distinct or necessary features of their artifacts by either direct communication by the participants during their final share out or directly observable from their artifacts. Those identified were intention, means, sensor, and body. Intention is the chosen purpose for their robot, or their initial idea. Means is denotes how the artifact fulfilled

their intention. Sensor refers to the means with which it senses the state of the world around it. Body is the overall physical embodiment of their final solution. Researchers experienced in working with children on robotics selected these features, and refined through discussion within the research team. Different populations or problems may require different attributes..

Attribute weights emphasize the importance of particularly difficult or design-critical features. Due to the pilot nature of this study and the field's lack of understanding about which parts of robotic engineering are particularly difficult for youth, we assigned all weights equally. Note that f_j is the weight for the attribute j, where $\sum_{j=1}^{n} f_j = 1.0$ (Shah et al., 2003). Since all of the weights must sum to one, we set all our weights to 0.25.

Ideas and features were mapped to attributes based upon observation of the artifact itself and discussion during a group's final share out. More specifically, we identified *intention* based on discussion during final share out, *means* both from discussion during share out as well as observing the artifact itself and both *sensor and body* through observing the artifact.

Values for the ideas were computed using the formula given in Shah et. al (2003): $S_j = \left[\left(T_j - C_j\right)/T_j\right] \times 10$ where T_j is the total number of ideas for attribute j and C_j is the number of instances for a specific idea in that attribute.

Finally, *Novelty scores are calculated* using a summation of those values, computed from $M = \sum_{j=1}^{n} f_j S_j$ (Shah et al., 2003). Resulting values range from 0 to 8.0 that were then translated to a percent of the possible value to get a Novelty Score out of 100.

In this case, the solutions for the LEGO group and the PaperBots group were scored as separate populations since although they were participating in the same activities; they were using different technologies with unknown difference in breadth of solution possibilities or impact on students' conversations.

Analyzing Collaborative Discourse

A coding scheme described by Barron (2003) was used to classify how students responded to a peer-proposed problem-solving solution. There is a two-part process for coding responses: identifying solution proposals and coding the responses as *Accept, Discuss* and *Non-engage*. A proposed solution was defined as any suggestion that explained how the group would help the designated character in the story. It was counted as an *accept* response if a group mate indicated "agreement with the content of the proposal," a *discuss* response if a group mate acknowledged "the proposal but did not accept them outright or reject them without rationale", and a *reject or ignore* response if a group member rejects "the proposal without a rationale….[or] there is a lack of relevant verbal response." The term *engage* refers to the both *accept* and *discuss* responses, and the term *non-engage* refers to the *reject or ignore* responses (Barron, 2003). In addition to the counts of the types of responses, an Engagement Score, Acceptance Score and a Discussion Score were calculated as the percent of *engage*, *accept* and *discuss* out of the total responses, respectively

One of the researchers coded all video data available for the Muncha Muncha Muncha. A second researcher randomly chose a group from the LEGO session and a group from the PaperBots session and independently coded their first hour of the activity using the same coding scheme. Agreement between the two coders was 100%.

Comparing Discourse and Solution Novelty

We used Microsoft Excel to calculate a Pearson product-moment correlation coefficient between each of the Engagement Score, Acceptance Score, and Discussion Score against Novelty Score.

Results

The methods described above were then applied to the data collected from the camp to yield the following results. Please note that all solutions were considered. The solutions in the book by the farmer were to dig a moat around his garden and after that failed to keep out the bunnies, build a large fortress wall around it. The bunnies were still able to infiltrate it through cunning though. The absurdity of the farmer's solutions and the anthropomorphic abilities of the bunnies opened up such possibilities that the student solutions, no matter how unrealistic, were considered as long as the students could present their reasoning which they all successfully did.

Solution Novelty

An example calculation: To clarify our process, we present an instance of mapping ideas to an attribute, computing the values for ideas and calculating a group's novelty score.

Using the LEGO session's artifacts shown in Figure 1 as an example, the *body* attribute had 3 different design concepts; a *single* body where all components were in a single package, a *single-functional* body where all components were in a single package but included some functional components like a cow catcher on the front, and a *tethered* system where the NXT brick was tethered to the functional portion of the robot.

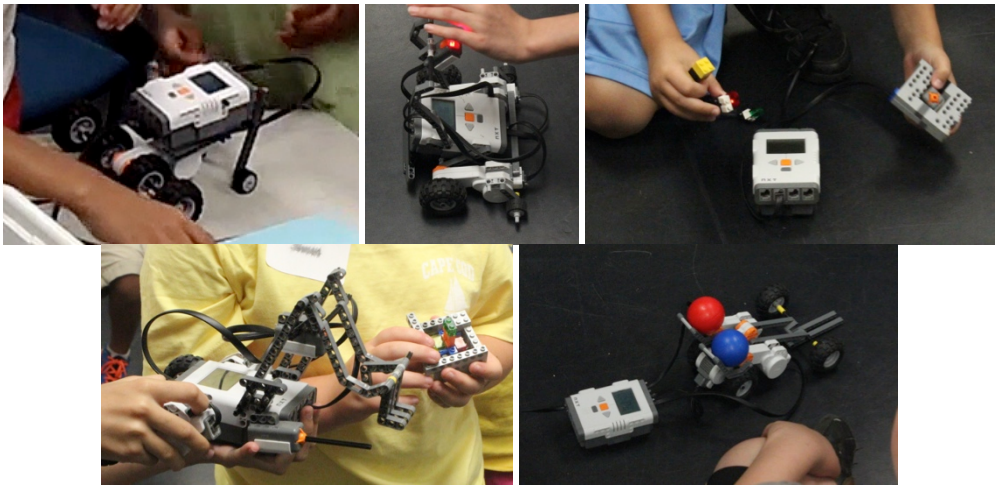


Figure 1. Artifacts from *Muncha Muncha Muncha* activity for LEGO group's A, B, and C in the top row and group's D and E across the bottom.

For the *single* and *single-functional* design ideas, there were two instances ($C_4 = 2$) of each so for a population of 5 ($T_4 = 5$) giving an idea value of $S_4 = [(T_4 - C_4)/T_4]\times 10 = [(5-2)/5]\times 10 = 6.0$. For the *tethered* idea, it was unique among that population ($C_4 = 1$) so it gains a higher novelty value of $S_4 = [(T_4 - C_4)/T_4]\times 10 = [(5-1)/5]\times 10 = 8.0$.

LEGO group C's novelty score comes from adding up all of their idea scores. They had the idea to *warn* the rabbits with a *light* triggered by a *pressure* sensor built into a *tethered* system robot. That would give them a value of 8.0, 8.0, 6.0, and 8.0 for their ideas (values shown in Table 1) and then, with the weights calculated in, a total novelty score of 7.2 (scores shown in Table 2).

Overall: We used the methods described above to compute the novelty of each group's design solution. Table 1 shows the results from computing the values for the ideas as described as step four. None of the attributes have all unique ideas, the intentions of the morning session being the least unique. Table 2 shows the ideas for each attribute by group as mapped in step three as well as their final novelty scores as calculated in step five. Overall, we found that the artifacts in the afternoon were on average more novel then the morning artifacts with statistical significance using a 1-tailed heteroscedastic t-test, which is statistically significant using the conventional 5% cut off.

Table 1: Attribute idea values from *Muncha Muncha Muncha* activity final artifacts and presentation

Lego Group											
Intention			Means			Sensing			Body		
Idea	C	V	Idea	C	V	Idea	C	V	Idea	C	V
Scare	4	2.0	Chase	3	4.0	User	1	8.0	Single	2	6.0
Help	1	8.0	Light	1	8.0	Dark	1	8.0	Sgl-Funct	2	6.0
			Catapult	1	8.0	Pressure	2	6.0	Tethered	1	8.0
						Baited	1	8.0			
PaperBots Group											
Intention			Means			Sensing			Body		
Idea	C	V	Idea	C	V	Idea	C	V	Idea	C	V
Scare	2	6.0	Scarecrow	2	6.0	None	1	8.0	Diorama	2	6.0
Trap	1	8.0	Drop Cage	1	8.0	Baited	1	8.0	Tethered	2	6.0
Help	2	6.0	Launch	1	8.0	Timed	1	8.0	Multi	1	8.0
			Pickaxe	1	8.0	Dark	2	6.0			

Discourse Markers for Collaboration

To illustrate the coding process, we provide two snippets of transcript and the correlated coding process. The first is from Group A, which had a lower Engagement Score. This could have been due to the fact the one girl,

Helen (gender-keeping pseudonyms have been used) seemed as if she did not want to work with her group mates so much as delegate tasks to them. Part of this is evident in the fact that George did not say anything during this active brainstorming session. This created a tension between Helen and Karl who seemed to want to involve George and work as a cohesive team. Despite this greater than normal tension, they are still productive in sharing and discussing ideas.

Karl: *drawing* So we could do- this is the whole vegetable garden. And then this is all the vegetables.
Helen: I have an idea. I could build a paper fence. And then-
Karl: No, because - you know how they tried that and it didn't work? We could do this-
Helen: No no no no no. I mean like a tall paper fence and then *looks at name tag* George- whatever your name is- you could fill the holes with something and then you *points to K* could build some kind of ceiling on it.
Karl: No like, yeah, I was thinking of the ceiling-
Helen: Yeah so you-
Karl: and like a door that you can open and close.
Helen: Um, the bunnies would be able to go through the door.

Table 2: Novelty scores from *Muncha Muncha Muncha* activity final artifacts and presentation

Lego Group	Intention	Means	Sensor	Body	Score	Novelty Score (out of 100)
	Wt = 0.25	Wt = 0.25	Wt = 0.25	Wt = 0.25		
A	Scare	Chase	User	Single	5	62.5
B	Scare	Chase	Dark	Single-Functional	5	62.5
C	Warn	Light	Pressure	Tethered	7.5	93.75
D	Scare	Chase	Baited	Single-Functional	5	62.5
E	Scare	Catapult	Pressure	Single	5.5	68.75
				Lego Avg. Score:	5.6	70
PaperBots Group	Intention	Means	Sensor	Body	Score	Novelty Score (out of 100)
	Wt = 0.25	Wt = 0.25	Wt = 0.25	Wt = 0.25		
Red	Scare	Scarecrow	None	Diorama	6.8	81.25
Blue	Trap	Drop cage	Baited	Tethered	7.8	93.75
Green	Help	Launch	Timed	Multi-functional	7.8	93.75
Yellow	Scare	Scarecrow	Dark	Tethered	6	75
White	Help	Pickaxe	Dark	Diorama	6.8	81.25
				PaperBots Avg. Score:	7.04	85

The solutions in this segment include the fence and the ceiling as they both explain how the group is keeping the bunnies out of the garden whereas the addition of the door does not explain how they are keeping the bunnies from the vegetables. Helen first introduces the idea of a fence on the second turn. Karl's response is categorized as *discuss* because even though he rejects the proposal, he explains that in the story, the character already tried building a fence and it did not prevent the bunnies from eating the vegetables. The idea of the ceiling is presented by Helen on the fourth turn and is immediately accepted by Karl so the response is coded as *accept*.

The second snippet is of group B which had a disruptive group member, Andrew, whom was constantly off task or suggesting inhumane solutions for keeping the bunnies out of the garden. Jane and Alexis who originally tried to include Andrew eventually started ignoring and rejecting his proposed solutions, possibly because they found them counter-productive or viewed them as his way of playing around. The two girls managed to complete the challenge with very limited assistance from Andrew.

Jane: No no no no no. Like, like a net! A net.
Andrew: No, its-
Alexis: But a net could hurt them.
Jane: No you just put it in it, like fish.

Andrew: Yeah, well what about like um like what about like a nuclear bomb? That won't hurt them.
Alexis: No
Jane: The bunny catcher.
...
Teacher: It's just going to drive around?
Alexis: It will go around- it will go around
Andrew: How is that suppose to scare them? Wait, no, it will drive around and then they exterminate them with like giant lasers-
Jane: If we could we would make it high speeds of running around
Teacher: Ok
Alexis: Yeah, we could scare the bunnies because a lot of animals if you get too close to them they get scared.
Teacher: They'll run away.
Alexis: Yeah, so we're just going to scare them away.
Andrew: These are radioactive bunnies

Table 3: Student Engagements from *Muncha Muncha Muncha* Activity

Group	Accept	Discuss	Non-engage	Total	Engagement Score (out of 100)	Acceptance Score (out of 100)	Discussion Score (out of 100)
A	5	3	7	15	53	33.33	20
B	2	6	4	12	67	16.67	50
C	3	6	1	10	90	30	60
D	2	2	2	6	67	33.33	33
E	N/A	N/A	N/A	N/A	N/A	N/A	N/A
LEGO	12	17	14	43	67	27	40
Red	2	2	1	5	80	40	40
Blue	4	10	2	16	87.5	27.78	62.5
Yellow	7	5	0	12	100	57.14	42
Green	1	7	2	10	80	10	70
White	2	3	2	7	71.43	28.57	42.86
Paperbots	16	27	7	49	86	32	54

The solutions proposed in this segment are a net, a nuclear bomb, and giant lasers. The idea of a net is introduced in the first turn by Jane and Alexis responds with concern for the bunnies being hurt by the net without accepting or rejecting the idea which means the response is categorized as *discuss*. A new proposal is brought to the group by Andrew at turn five for a nuclear bomb where he assures his group mates that it would not hurt the bunnies, however he is flat out rejected by Alexis which is coded as *non-engage*. Later, when Jane and Alexis are explaining to the teacher their current solution of a robot that drives around and scares bunnies, Andrew proposes giant lasers. This idea is not addressed in six turns, thus it is considered being ignored and categorized as *non-engage*. The results of the discourse coding are summarized in Table 3.

Collaborative Discourse Predicts Solution Novelty

Each group's Novelty Score was charted and regressed against their Engagement Score (Figure 2), Acceptance Score and Discussion Score. Because one video file was lost, we only had enough video to code nine groups.

Our regression value (R) of .90 for *Engage* responses is statistically significant ($p < 0.00042$), indicating a strong correlation. However, when Accept and Discuss scores were charted against Novelty score, there was not a significant correlation ($p < .13$ and .08 respectively), indicated by the low regression values of 0.40 and 0.50.

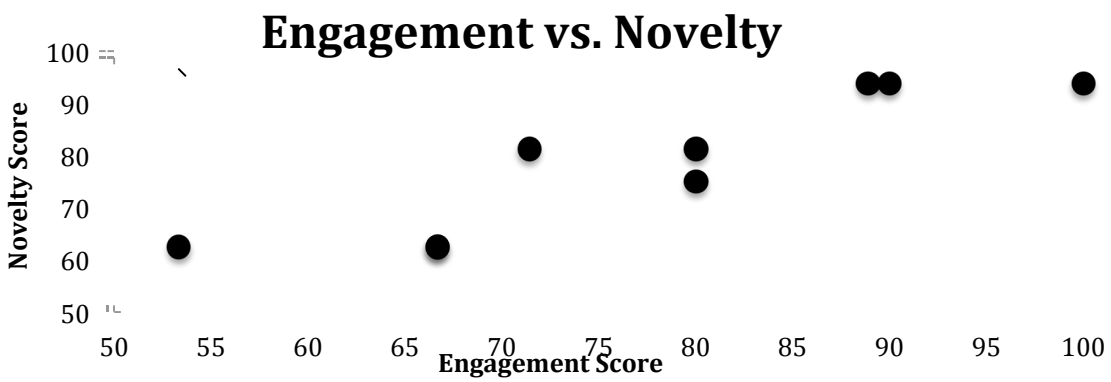


Figure 2. Engagement versus novelty

Discussion

Our results show that there is a strong correlation between number of *engage* responses and novelty in the students' final artifacts. Engaged responses in student discourse predict novel solutions. There has been other research that shows there are ways to teach this type of discourse to students (Demetriadis, Egerter, Hanisch, & Fischer, 2011). Further, we have demonstrated the possibility of novelty assessment being used in a classroom-like environment using only observations, share outs and pictures of final artifacts.

By identifying a type of student discourse that supports novelty, we began research that we hope will ultimately inform teachers how to better foster creativity in the classroom. The method utilized in this paper is a quantitative way of defining the type of student discourse but in classrooms, in the moment, discourse must be observed and the most beneficial discourse be fostered. "It's clear that the classroom teacher plays a critical role in establishing and modeling practices of productive group learning processes and conversations. Observing a group's interactions can provide teachers with valuable insight" (Barron & Darling-Hammond, 2008) but only if they know what they are looking for. CSCL research has explored how a facilitator can influence discourse. One example is having a tutor sustain and deepen inquiry through well-timed refocusing (Lakkala, Muukkonen, & Hakkarainen, 2005). It has long been taught that group members shouldn't say "types of comments that indicate competition, premature judgment, or failure to listen in group discussion" (McKendall, 2000). Our research supports this long-held idea and extends it to teachers actively supporting engaged responses in student discourse.

This study also illustrates the use of a creativity assessment instrument, specifically the method of assessing novelty described by Shah et. al (2003), in a classroom-like environment. The prescribed novelty assessment allows teachers to grade projects after-the-fact either by looking at the physical artifacts, pictures or video, depending on their preferences and resources. This is a vast improvement over other instruments that advocate identifying the provenance of an idea, not something that can readily be applied in a classroom, though advancements in learning analytics may change this (Blikstein et al., 2012). The methods used in their current form are unrealistic for timely use in classrooms but the novelty measure and others like it in the works of Shah et al (2003) may be useful for assessment as part of an application that organizes the student works and their features and takes care of the calculations. Despite the small sample size, the implications of a quick and easy way to measure even an aspect of creativity after-the-fact are clear. Teachers who have been able to intuitively tell that one solution is more original than others will finally have a way to measure and support that sense. In the future, systematic assessment creativity could even become a more effective way to measure teamwork effectiveness than current methods of asking groups to report on the other members or scattered observations.

Conclusions

Open-ended and group problem solving have been shown to lead to more robust individual understandings (Kapur, 2008). While teachers are constantly pressured to increase test scores, they are also expected to promote teamwork and creativity through engineering design projects (NGSS, 2013). This study suggests that it may be possible to combine methods from the learning sciences, art, and engineering education research to better analyze creative problem solving, and ultimately to develop classroom-practicable techniques for assessing creativity.

References

Barron, B., & Darling-Hammond, L. (2008). Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning. *Powerful learning: What we know about teaching for understanding*, 11-70.

Barron, B. (2003). When smart groups fail. *The journal of the learning sciences*,*12*(3), 307-359.
Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education—A systematic literature review. *Thinking Skills and Creativity*, *8*, 80–91.
Demetriadis, S., Egerter, T., Hanisch, F., & Fischer, F. (2011). Peer review-based scripted collaboration to support domain-specific and domain-general knowledge acquisition in computer science. *Computer Science Education*, *21*(1), 29–56.
Develaki, M. (2010). Integrating Scientific Methods and Knowledge into the Teaching of Newton's Theory of Gravitation: An Instructional Sequence for Teachers' and Students' Nature of Science Education. *Science & Education*, *21*(6), 853–879.
Dillenbourg, P. (1999). What do you mean by collaborative learning?.*Collaborative-learning: Cognitive and computational approaches.*, 1-19.
Eteläpelto, A., & Lahti, J. (2008). The resources and obstacles of creative collaboration in a long-term learning community. *Thinking Skills and Creativity*, *3*(3), 226–240.
Glück, J., Ernst, R., & Unger, F. (2002). How Creatives Define Creativity : Definitions Reflect Different Types of Creativity. *Creativity Research Journal*, *14*(1), 55–67.
Kapur, M. (2008). Productive Failure. *Cognition and Instruction*, *26*(3), 379–424.
Kurtzberg, T. R., & Amabile, T. M. (2001). From Guilford to creative synergy: Opening the black box of team-level creativity. *Creativity Research Journal*,*13*(3-4), 285-294.
Lakkala, M., Muukkonen, H., & Hakkarainen, K. (2005). Patterns of scaffolding in computer mediated collaborative inquiry. *Mentoring & Tutoring: Partnership in Learning*, *13*(2), 281–300.
McKendall, M. (2000). Teaching Groups to Become Teams. *Journal of Education for Business*, *75*(5), 277–282.
Merrill, J. A., Charyton, C., & Jagacinski, R. J. (2008). CEDA: A research instrument for creative engineering design assessment. *Psychology of Aesthetics, Creativity, and the Arts*.
NGSS. (2013). *Engineering Design in the NGSS* (pp. 1–7).
O'Connell, B. (2013). PaperBots: An Inexpensive Means for Engineering Education. In *120th ASEE Annual Conference & Exposition*. American Society for Engineering Education.
Peterson, C. W., & Madsen, C. K. (2010). Encouraging Cognitive Connections and Creativity in the Music Classroom. *Music Educators Journal*, *97*(2), 25–29.
Renzulli, J. S. (2005). The Three-Ring Conception of Giftedness: A developmental model for promoting creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 246–279). Cambridge University Press.
Saunders, R., & Gero, J. (2002). How to study artificial creativity. *Proceedings of the 4th conference on Creativity & ...*, 80–87.
Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE Project: Trying to Bring the Classroom into World 3. In *Classroom lessons Integrating cognitive theory and classroom practice* (pp. 201–228).
Shah, J. J., Vargas-hernandez, N., & Smith, S. M. (2003). Metrics for measuring ideation effectiveness. *Design Studies*, *24*, 111–134.
Soller, A., & Lesgold, A. (2007). Modeling the process of collaborative learning. In *The Role of Technology in CSCL* (pp. 63-86). Springer US.
Sternberg, R. J. (2005). WICS: A Model of Positive Educational Leadership Comprising Wisdom, Intelligence, and Creativity Synthesized. *Educational Psychology Review*.
Sullivan, F. R. (2011). Serious and Playful Inquiry : Epistemological Aspects of Collaborative Creativity, *14*, 55–65.
Taylor, C. W. (1988). Various approaches to and definitions of creativity. *The nature of creativity*, 99-121.
Van Harpen, X. Y., & Presmeg, N. C. (2013). An investigation of relationships between students' mathematical problem-posing abilities and their mathematical content knowledge. *Educational Studies in Mathematics*, *83*(1), 117–132.
Webb, A. N., & Rule, A. C. (2012). Developing Second Graders' Creativity Through Literacy-Science Integrated Lessons on Lifecycles. *Early Childhood Education Journal*, *40*(6), 379–385.
White, T. (2006). Code talk: Student discourse and participation with networked handhelds. *International Journal of Computer-Supported Collaborative Learning*.
Yuan, X., & Sriraman, B. (2011). An exploratory study of relationships between students' creativity and mathematical problem-posing abilities. In *The elements of creativity and giftedness in mathematics* (pp. 5-28). SensePublishers.

Acknowledgements

This project is funded by the National Science Foundation DRK-12 program, grant # DRL-1020243, as well as by LEGO. Any opinion, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF or LEGO. We thank Merredith Portsmore, Elissa Milto, Dan Wise, Joe Sanford, and Chelsea Andrews for their assistance.

Supporting Pre-Service Science Teachers' Planning of Task-Based Classroom Discussions

Danielle Ross, Aaron M. Kessler and Jennifer Cartier, University of Pittsburgh, 5500 Wesley W. Posvar Hall
Email: dross1225@gmail.com, aaronmkessler@gmail.com, jcartier@pitt.edu

Abstract: This study focuses on a set of core, or high-leverage, practices in teacher education in order to allow pre-service teachers to begin to develop a set of necessary skills to successfully support student learning through inquiry (Grossman et al., 2009b). Specifically, in the secondary science teacher preparation program teacher educators adopted a practice-based focus in which secondary pre-service science teachers participated in the high-leverage practice of engaging students in task-based science discussions and the planning of those discussions. The results suggest that pre-service teachers are able to incorporate aspects of these instructional models in their discussion planning practices. The results also suggest a shift in focus on different aspects of planning over time and across science disciplines that open the possibilities for future work.

Background and Significance of Work

Teachers will face many challenges as we move forward into the age of the *Next Generation Science Standards (NGSS)* (Achieve, 2013). The *NGSS* aim to develop a population of scientifically literate and talented students who can participate in the "innovation-driven economy" (p. 1). In order to meet these goals, teachers must provide students with opportunities to engage in science and engineering practices (SEPs), e.g., interpreting data, engaging in argument, and learning core ideas of these disciplines.

In classrooms where students are engaged in the practices of scientists and engineers, teachers face the additional challenge of designing instruction so that students have opportunities to wrestle with the underlying science ideas at a high level (Engle, 2011; Engle & Conant, 2002; Smith & Stein, 2011). To design instruction in this way, a teacher must first identify key learning goals to focus the lesson and then choose a task that is robust enough to support students' thinking and learning in the discipline. After selecting (or designing) a task, the teacher must then imagine in detail the ways in which her students might engage with the task, design appropriate tools and scaffolds to support and direct that engagement, and plan for ways to monitor students' work during the task, also know as the Five Practices model (Cartier, Smith, Stein, & Ross, 2013). We assert that this type of planning is both sophisticated and teachable. In this study, we describe our approach to helping pre-service secondary science teachers (PSTs) adopt effective (high-leverage) planning practices. We investigate the impact of engaging PSTs in a series of carefully scaffolded exercises focused on helping them to learn how to anticipate and support students' engagement in challenging science tasks and subsequent discussions about those tasks. Specifically we examined:

(1) To what extent do PSTs use the Five Practices instructional model in their planning?
(2) In what ways do PSTs' use of the Five Practices instructional model change over time?

Importance of Classroom Discourse in Science

Students in today's science classrooms must have opportunities to develop the practices and skills used in science and engineering professions in order to be productive members of our technologically advanced society (Achieve, Inc., 2013; Duschl, 2008). Discourse – or students engaging in talk with one another around disciplinary concepts – is a key component of classrooms where students are engaged productively in science and engineering practices (SEPs). While discourse is necessary to achieve the *NGSS* goals, it is also a challenge for teachers to orchestrate (Grossman et al., 2009a; Stein, Engle, Smith, & Hughes, 2008).

As teacher educators, our goal is to provide novice teachers with conceptual and practical tools to support their learning and teaching (Grossman, Hammerness, & McDonald, 2009b), and we are particularly interested in supporting the practices related to orchestrating productive classroom discussion. Researchers have identified many different pedagogical strategies designed to aide teachers in supporting robust discussions and supporting students in the types of discourse that increase deep understanding. Pedagogical frameworks, such as Investigating and Questioning our World Through Science and Technology (IQWST) (Berland & Reiser 2008; McNeill, Lizotte, Krajcik, & Marx, 2006), the evaluate-alternatives model (Sampson & Grooms, 2009), the Accountable Talk framework (Michaels, O'Connor, & Resnick, 2008), and the Five Practices model (Cartier, Smith, Stein & Ross, 2013; Smith & Stein, 2011; Stein et al., 2008) present teachers with micro-practices that support student learning through discussion. These frameworks have several features in common. Specifically, each emphasizes the need for teachers to (1) choose appropriate instructional content that promotes discourse, (2) guide and support students through scaffolding, and (3) hold students accountable to classroom and scientific norms.

Expert teachers have established routines and plans that help to support such ambitious instruction. Many of their moves have been developed and practiced over time and are often not taught to beginning teachers (Leinhardt & Steele, 2005). However, we assert that because novice teachers struggle to support productive student engagement in science discussion, and because this type of instruction shows so much promise for enabling learners to engage in SEPs and master core content, teacher educators should provide pre-service teachers with conceptual and practical tools to support this practice (Grossman, Hammerness, & McDonald, 2009). Consequently, we have designed pedagogy courses for PSTs that include explicit opportunities to learn about and practice elements of the Five Practices instructional model (Cartier et al., 2013; Smith & Stein, 2011). By supporting teachers in learning how to *anticipate* student thinking, *monitor* student responses to tasks, *select* students to present their work, purposefully *sequence* these presentations, and *connect* the ideas through discussion, this model guides teachers through the processes of preparing for and supporting whole class discussions (Cartier et al., 2013; Smith & Stein, 2011). We leveraged these five practices in the design of the secondary science methods course.

Theoretical and Methodological Approaches

In order to make PSTs more comfortable with enacting discussions in their classrooms, we chose to attend to specific aspects of this practice and designed learning experiences to help them develop the skills to enact productive discussions in their own 7^{th}-12^{th} grade classrooms. We utilized the framework developed by Grossman and colleagues (2009a), to design cycles of decomposition, representation, and approximation of planning practices related to task-based discussions. By separating complex practices in subsets, or micro-practices, such as the critique and analysis of lesson planning and discussion facilitation, into its component parts, we reasoned that PSTs would feel more comfortable enacting the practices in their own classrooms, leading to consistent implementation of these practices over time (Grossman et al., 2009a; Stein et al., 2008). What follows is a detailed example of how we utilized this framework within the context of the pedagogy courses for secondary science PSTs at the Midwestern university where this study occurred (see Table 1).

We divided the fall semester into lessons in which the PSTs engage in science as learners and as practitioners through iterative cycles of decompositions, representations, and approximations of practice (Grossman et al., 2009a). More specifically, in addition to other practices, the PSTs observed teacher educators decompose and represent the practice of the Five Practices model by having the PSTs approximate components of this model by planning discussion-based lessons using this model, and rehearsing and formally teaching instructional episodes with peers. Through varying levels of authenticity the teacher educators guided the PSTs to examine specific planning and instructional practices and certain teacher moves that help to support discussion orchestration. Once the PSTs had the opportunity to unpack these instructional micro-practices, they were better able to use them in their own planning and teaching (Grossman et al., 2009a).

Decomposition

Enacting productive task-based discussions in any discipline is a complex practice in which teachers employ a variety of moves, micro-practices, and routines during instruction (Leinhardt & Steele, 2005). In order for PSTs to begin to engage in this or any practice, Grossman et al (2009b) posit that PSTs may need varying opportunities to recognize and then enact small components of professional practice after which they can then begin to integrate each micro-practice completely.

By decomposing the Five Practices model into its distinct practices, we provided PSTs with an opportunity to focus on certain fundamental skills, e.g. anticipating student thinking, that will help them to prepare for and facilitate productive science discussions (Grossman et al., 2009a). This decomposition allowed the instructor to call attention to as well as provide immediate feedback to students as they analyzed and reflected upon these components. Through this feedback, the PSTs attend to particular teacher moves and aspects of the instructional model that help support their discussion planning and facilitation. By focused attention on certain aspects of student thinking, student work, and important teacher moves, these aspects of discussion typically viewed as improvisational by many beginning teachers seem less so (Smith & Stein, 2011; Stein et al., 2008). By giving the PSTs this instructional planning tool and facilitating a discussion, our goal was to help them feel more comfortable standing to the side of the dialogue and allowing students' opportunities to engage with each other.

In order to support the development of the PSTs' planning for lessons where students engage in science discussions, we selected particular practices based on past research (Smith & Stein, 2011), namely: writing specific learning goals, anticipating student thinking, planning for monitoring, selecting and sequencing student approaches, and connecting student ideas and disciplinary ideas. We believe that the development of these practices in PSTs' repertoires will best support the development of their capacity to design task-based science discussions.

Representation

The PSTs observed expert teachers utilizing the model, read written cases, as well as analyzed evidence of student work. Through varying levels of authenticity we guided the PSTs to examine specific practices or certain teacher moves that help to support the instructional dialogue that might otherwise go unnoticed. By drawing attention to particular details, the PSTs began to notice and learn ways in which they might begin to build their own teaching repertoire.

Once the PSTs analyzed these micro-practices related to planning science discussions, they have a model, or representation, by which to analyze this complex practice (Stein et al., 2008). For example, by providing PSTs with examples of student work and a case study of how a classroom teacher implements her classroom discussion, we foregrounded salient aspects of anticipation, monitoring, selecting, sequencing, or connecting the teacher used. Through these various representations, we supported the PSTs in visualizing ways in which they can begin to use and develop their own identity as a classroom teacher (Grossman et al., 2009a).

Approximation

By simulating and role-playing a Five Practices discussion in the methods classroom, the PSTs engaged in approximations of practice similar to those identified by Grossman et al (2009a). As they gained experience, they engaged in varying levels and iterations of authentic and complex discussion practices, thereby developing the knowledge and skills necessary to begin to integrate the decomposed pieces of the model. Through this public practice and feedback provided to the PSTs, we highlight particular aspects of the model, while other, less important, aspects of discussion facilitation are ignored. By drawing PSTs' attention to these important aspects and allowing them to engage in opportunities to practice, they can begin to develop their PDC to design tasks necessary to facilitate productive, engaging science discussions with students.

In designing the secondary science methods courses, we adopted Ball and Forzani's (2009) practice-based focus to design a Teaching Laboratory course for secondary science pre-service teachers in which they participate in the high-level practice of supporting productive classroom discourse. We designed role-play scenarios that enabled our PSTs to engage in approximations (Grossman et al., 2009a) of these Five Practices in the context of the Teaching Laboratory course. Here, we provide examples of our approximation scenarios and described the principles underlying the design (see Figure 1).

Grossman et al.'s framework Component	*Description*	*Example from Teaching Lab course*
Decomposing the Practice	Breaking down the overall practice into component parts is necessary in order for the practice itself to be accessible to learners. **Thus, a complex practice like "orchestrating classroom discussion" can be viewed as consisting of several smaller practices.**	Smith and Stein (2011) decompose the practice of orchestrating discussion into five sub-practices: (1) anticipating students' ideas and/or problem solutions related to an instructional task; (2) monitoring students' work during the task and taking note of which ideas emerge; (3) selecting examples of students' work to highlight in whole-class discussion; (4) planning the sequence in which you will discuss these work samples; and (5) planning questions and talk that will elicit key ideas and help learners connect their ideas to one another and underlying canonical disciplinary ideas.
Representing the Practice	Representations should be chosen or created to highlight certain aspects of practice while backgrounding others. Representations may be teaching artifacts like lesson plans or excerpts from curriculum materials, or they may be videos, transcripts, or examples of student work.	In our work, we developed "sample stories" that are partial transcripts of classroom discussions. These transcripts are accompanied by meta-language that reveals the teacher's rationale for decision-making during the discussion.
Approximating the Practice	Learners develop facility with a practice only through participation in approximations of that practice. **Initial approximations are usually constrained and simplified.** Later, learners encounter the practices in more complex and realistic contexts.	We developed elaborate role-play scenarios related to various science ideas (e.g. kinetic molecular theory, the theory of natural selection, mathematical modeling of radioactive decay). The materials that support each role-play scenario include: (a) a description of the instructional activities in which students would participate; (b) samples of student work that have been selected or invented such that typical alternative conceptions are represented; (c) background information for the person playing each student's role; and (d) tools to support teachers' monitoring, selecting, sequencing, and question planning. The role-play scenarios provide opportunities for pre-service teachers to engage in approximations of all five sub-practices related to orchestrating discussion. They take turns adopting the role of student and teacher throughout the scenarios and have multiple opportunities to offer and receive feedback on their teaching performance and decision-making throughout the each scenario.

Figure 1. Description of the Approximation Scenarios From the Teaching Laboratory Course (Ross, Cartier, Forman, & Walker, 2011).

Participants

A total of 18 teachers enrolled in the 2011-2012 secondary science Master of Arts in Teaching and professional year teacher certificate year program. The 14 subjects (5 males and 9 females) of this study were a sample of convenience because they were all enrolled in the science methods course and required to participate in the coursework. The majority were 22-26 years old and Caucasian. Two instructors taught the Teaching

Laboratory during this time using a common syllabus and identical classroom learning tasks and course assignments.

Data Corpus and Coding

Over the course of 10 weeks during the fall 2011 term, the PSTs completed two instructional planning (IP) assignments in which they planned and designed their own lessons focusing on whole class discussion around data (see Figure 2). Each PST implemented their lesson at the secondary science placement in which they were assigned. Upon completion of the lesson, PSTs completed a reflection on the lesson and posted all relevant materials (e.g. lesson plan, tasks, reflection, etc.) to a shared online planning tool that the research group had access. In service of this study's goals we focused the analysis specifically on the PSTs planning documents.

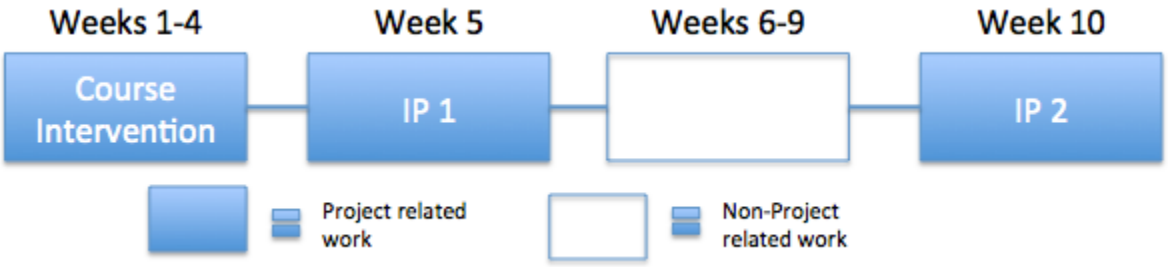


Figure 2. Intervention and instructional planning timeline.

Recall that the purpose of this study is to understand what features of the Five Practices (Stein & Smith, 2011) the PSTs, who had participated in the role-play scenarios around these practices, utilize when planning for a discussion lesson. We coded each lesson plan for instances of anticipating, monitoring, selecting, sequencing, and connecting as defined by Stein and Smith (2011). Twenty-five percent of the data set was double coded with an interrater reliability of 85%. In areas where there was disagreement, we discussed the disagreements and reached consensus.

Major Findings

Use of the Five Practices Model in Planning

In both lesson plans, PSTs showed evidence of using the Five Practices in their lesson planning. Coding for use of the Five Practices included a variety of different ways that the PSTs used this instructional model. PSTs anticipated, monitored, selected, sequenced, and connected in their planning. Exemplars of each of these practices are below (Table 2).

Table 2. Examples of Use of the Five Practices in the PSTs' Lesson Planning.

Anticipating	"Based on the surveys I had students complete in the last class, I have already organized them into groups specific groups with similar thinking patterns (though not 100% the same, because differences will help to facilitate a discussion)." Details misconceptions and groups and what their underlying misconceptions are for each group."
Monitoring	Plans questions and prompts to address to students and probe their thinking in order to generate a discussion. Creates a completed monitoring tool based on student work from previous day.
Selecting	"I will ask for volunteers if students fail to volunteer I will ask for group 2 to start." They had the q2m2 then group 5 can offer alternative explanation of lab data then group 1 with the alternative form of sugar (I will privately ask Brittney from group 4 if she would come up she will wrap up this discussion b/c she had most complete and well thought out response"
Sequencing	"I will ask for volunteers if students fail to volunteer I will ask for group 2 to start." They had the q2m2 then group 5 can offer alternative explanation of lab data then group 1 with the alternative form of sugar (I will privately ask Brittney from group 4 if she would come up she will wrap up this discussion b/c she had most complete and well thought out response"
Connecting	Plans questions that prompt students to make comparisons to the affordances and drawbacks of sickle cell anemia.

In the first set of lesson plans (Instructional Performance 1), coding for use of the Five Practices model showed that PSTs did utilize part of the model in their planning. Of the 14 PSTs, evidence of anticipating and monitoring occurred the most in their planning practices with 8 out of 14 anticipating and 11 out of 14 monitoring to some degree (see Table 3). Half of the teachers (7) showed evidence of selecting students' work to present in their planning, with sequencing and connecting used by only 5 teachers.

In Instructional Performance 2 (IP 2) planning, PSTs show evidence of using connecting in their planning most frequently, 10 out of 14. Selecting and sequencing use occurred the least in their planning, 4 and 2 PSTs, respectively. However, use of anticipating and monitoring practices decreased compared with IP1, 6 PSTs and 7 PSTs respectively (see Table 1). Overall, there is not evidence that the PSTs use the Five Practices model as intended, incorporating all of the Practices in their planning. Instead, PSTs show evidence of using selective practices in each lesson plan

Table 3: Total number of pre-service teachers' use of features of the five practices instructional model in their lesson planning assignments.

	Anticipating	Monitoring	Selecting	Sequencing	Connecting
IP1					
No Evidence	6	3	7	9	9
5P Use	8	11	7	5	5
IP2					
No Evidence	8	7	10	12	4
5P Use	6	7	4	2	10

Use of the Five Practices Model in Discussion Planning Over Time

In looking at the way PSTs planning changed over time, we see two patterns emerge. First, there is a shift in anticipating and monitoring use in IP1 with minimal attention paid to planning for connecting to clear evidence in IP2 where connecting is a focus of the PSTs' planning attention (see Table 1). This finding suggests that PSTs recognized the importance of planning for connecting and discussion orchestration after implementing IP1 and made that planning a focus in their subsequent lesson plans. Moreover, evidence suggests that PSTs attend to particular aspects of their planning at certain times. For example, immediately after coursework focusing on the importance of using the Five Practices, 51% of the PSTs used elements of the Five Practices model in their planning. As time passes, only 41% of the PSTs used elements of the Five Practices model in their planning. This finding suggests: (1) PSTs recognized the importance of planning for connecting in their discussion, (2) PSTs focused on other elements of their planning as they continue teaching in their field placements.

Second, in an effort to determine if any content specific variations occurred in the PSTs planning, we analyzed the data with respect to content area (see Tables 4 and 5). Although the sample size is low (n=14), a pattern emerges between Physical Science and Biological Science. For IP 1, both content area PSTs plan using features of the Five Practices at some level. However, for IP 2, there is a significant drop in the General Science PSTs that plan using the Five Practices model. The Biological Science PSTs plan using the Five Practices across IP 1 and IP 2, while planning in the Physical Sciences drops noticeably. This finding indicates a possible connection between planning and using the Five Practices model in the two science areas. We cannot speak to specifics without further research, but there may be a connection between content area and curriculum and ease or feasibility of using the Five Practices model. Moreover, there may be differences in the way the PSTs understanding of the Nature of Science or learning of science. For instance, if a PST has an understanding that Chemistry is learned through doing experiments and completing problems, then it might be difficult to conceptualize a robust whole class discussion around Chemistry content.

Table 4: Total number of pre-service teachers' use of the five practices model in planning for instructional performance 2 by content area.

		IP1				
		Anticipating	**Monitoring**	**Selecting**	**Sequencing**	**Connecting**
Physical Science	No Evidence	4	2	5	6	7
	5P Use	6	8	5	4	3
	Total	10	10	10	10	10
Biological Science	No Evidence	2	1	2	3	2
	5P Use	2	3	2	1	2
	Total	4	4	4	4	4

Table 5: Total number of pre-service teachers' use of the five practices model in planning for instructional performance 3 by content area.

		IP2				
		Anticipating	**Monitoring**	**Selecting**	**Sequencing**	**Connecting**
Physical Science	No Evidence	7	6	8	9	4
	5P Use	3	4	2	1	6
	Total	10	10	10	10	10
Biological Science	No Evidence	1	1	3	3	0
	5P Use	3	3	1	1	4
	Total	4	4	4	4	4

Conclusions and Implications

These findings suggest that the Five Practices role-play approximations in the Teaching Laboratory course plays a role in supporting PSTs as they plan for robust science discussions using this instructional model. Over time, the PSTs appear to focus on other aspects of their teaching, e.g., classroom management, and less focused on planning related specifically to discussions using the Five Practices model. It might be necessary to design supports methods courses to further unpack the importance of student talk and discussion in science. In addition, we are currently designing ways to support PSTs planning and teaching through discussion throughout their teacher preparation coursework. This study provides evidence that the Five Practices model and teaching of this instructional model provides a useful framework for PSTs as they plan discussions. By simulating and approximating a Five Practices discussion in pedagogy courses, the PSTs engaged in approximations of practice similar to those identified by Grossman et al (2009a) as effective. As the teachers gained experience, they engaged in varying levels of authentic and complex discussion practices, thereby developing the knowledge and skills necessary to begin to integrate the decomposed pieces of the model as they planned lessons for their own 7-12 students. Through the public practice and feedback, we highlighted particular aspects of the Five Practices important for planning a productive discussion. By drawing PSTs' attention to these important aspects and allowing them to engage in opportunities to practice, they showed evidence of using these practices as they planned to engage their students in science discussions.

Additionally, these findings suggest that the role-play scenarios in which the PSTs participated in the pedagogy courses are effective in the PSTs learning. Providing the PSTs with opportunities to engage in the approximations of practice (Grossman et al., 2009) allows for greater traction and understanding of the strategies presented in the role-play. Problematizing the issues that arise as teachers enact discussions in the familiar and comfortable setting of their own classrooms allows the PSTs to notice these issues and practice ways to address them.

In addition, findings from this study indicate the importance of grounding Five Practices discussions in high-levels tasks (Smith & Stein, 2011). Although not a focus of this study, anecdotal data reveals that designing and planning discussions around tasks that are not cognitively demanding for students can hinder the ways in which the PSTs plan and implement these discussions. As such, we continue to work on designing ways in teacher preparation to support PSTs development as instructional designers.

In order for the ambitious vision of science instruction presented by the *NGSS* (Achieve, Inc., 2013) to become a reality in secondary schools, teachers must design instruction with these goals in mind. By using various curriculum resources (e.g. texts, online lesson plans and resources, standards, curriculum materials, etc.), teachers can design instruction that supports students' engagement in SEPs and their sense making related to key disciplinary phenomena. The ability to navigate through the vast number of these resources and to design instruction appropriate for each group of students is the essence of what Brown (2009) terms *pedagogical design capacity (PDC)*. Thinking about the task, or thinking through the lesson, in critical ways can support teachers' as they design their planned curriculum (Smith & Stein). Analyzing tasks through the student lens allows teachers to understand their value in supporting students' learning and support teachers as they build upon individual tasks to create lesson arcs and curriculum units (Remillard, 2000).

Theoretically, we will continue to learn more about teacher learning and the mechanisms behind learning to orchestrate discussions through the continued use of the Five Practices model and the iterations of representation, decomposition, and approximation of practice through design research that continues in these courses (Grossman et al., 2009a). The potential learning ecologies and theories that emerge from this research will enhance teacher education across disciplines. Practically, it directly impacts the secondary science pre-service teachers in this teacher preparation program yearly. Ultimately, teachers impact secondary science student learning greatly through participating in task-based discussions regularly. This study lays the foundation for generalizability to other disciplines. The features of the role-play instructional model based on the Five Practices (Stein et al., 2008) can be used in teacher education course in all disciplines. Through the

identification of important features of the model that are essential for PST learning, teacher educators can design instruction to incorporate similar role-play instructional models in their disciplines.

The work associated with preparing pre-service science teachers to incorporate science and engineering practices into their instruction is something everyone should be concerned with, especially given the release of the *Next Generation Science Standards*. As the demands on teachers' instruction increases, beginning teachers need to be able to incorporate high-leverage practices effectively in order to engage students. This work describes a feasible way to prepare PSTs to engage students in task-based science discussions and the planning of those discussions through a set of decompositions, representations, and approximations. Additionally, this work addresses what these practices can look like in real environments, suggests one possible way to frame these practices in established instructional theory, opens the doors for discussions around practices, and has some clear avenues for further research including lesson implementation and student learning.

References

Achieve, Inc. (2013). *The Next Generation Science Standards*.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is – or might be – the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*, 6 - 8, 14.

Ball, D. L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education, 60*(5), 497-511.

Ball, D. L., Sleep, L., Boesrt, T. A., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *The Elementary School Journal*, 109 (5), 458-474.

Berland, L. K., McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94, 765-793.

Berland, L. K., & Reiser, B. (2008). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

Beyer, C. & Davis, E. A. (2009). Supporting pre-service elementary teachers' critique and adaptation of science lesson plans using educative curriculum materials. *Journal of Science Teacher Education* 20, 517-536.

Beyer, C., Delgado, C., Davis, E., & Krajcik, J. (2009). Investigating teacher learning supports in high school biology curricular programs to inform the design of educative curriculum materials. *Journal of Research in Science Teaching 46*(9), 977-998.

Brown, M. W. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eisenmann, and G. M. Lloyd (Eds.), *Mathematics Teachers At Work: Connecting Curriculum Materials and Classroom Instructio*n. New York, NY: Routledge, Taylor and Francis.

Cartier, J. L., Smith, M. S., Stein, M. K., & Ross, D. K. (2013). *5 Practices for Orchestrating Productive Science Discussions*. Reston, VA: National Council of Teachers of Mathematics and Corwin Press.

Cazden, C. B. (2001). *Classroom Discourse: The Language of Teaching and Learning (2nd ed.)* Portsmouth, NH: Heinemann.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9–13.

Davis, E. A., & Smithey, J. (2009). Beginning teachers moving toward effective elementary science teaching. *Science Education*, 93(4), 745-770.

Doyle, W. (1983). Academic work. *Review of Educational Research,* 53(2): 159-199.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education* (pp. 268-291). London: Sage Publications.

Engle, R. A. (2011). The Productive disciplinary engagement framework: origins, key concepts and developments. In David Yun Dai (Ed.) *Design Research on Learning and Thinking in Educational Settings: Enhancing Intellectual Growth and Functioning*, 161-200. London: Taylor & Francis.

Engle, R. A., & Conant, F. C. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners in a classroom. *Cognition and Instruction*, 20(4), 399-483.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record* 103(6), 1013-1055.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009a). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055-2100.

Grossman, P., Hammerness, K., & McDonald, M. (2009b). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: theory and practice*, 15(2), 273-289.

Grossman, P., & Thompson, C. (2008). Learning from curriculum materials: Scaffolds for new teachers? *Teaching and Teacher Education*, 24, 2014-2026..

Leinhardt, G., & Steele, M.D. (2005). Seeing the complexity of standing to the side: Instructional dialogues. *Cognition and Instruction*, 23(1), 87-163..

McNeill, K. L., Lizotte, D. J., Krajcik, J. & Marx, R. W. (2006). Supporting students construction of scientific explanations be fading scaffolds in instructional materials. *Journal of the Learning Sciences*, 15(1), 153-191.

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and civic life. *Studies in Philosophy and Education*, 27, 283-297.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). Sage Publications, Inc.

Remillard, J. T. (2000). Can curriculum materials support teachers' learning? Two fourth-grade teachers' use of a new mathematics text. Elementary School Journal, 100(4), 331–350.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211-246.

Ross, D. K., Cartier, J. L., Forman, E., & Walker, D. C. F., (2011). Using Role Play Scenarios to Help Pre-Service teachers Develop Foundational Discursive Practices. Poster Presentation at the *Science and Mathematics Teacher Imperative,* Portland, Oregon.

Sampson, V. & Grooms, J. (2009). Promoting and supporting scientific argumentation in the classroom: The evaluate-alternatives instructional model. *Science Scope*, 33.1, 66.

Smith, M. S., & Stein, M. K. (2011). *5 Practices for Orchestrating Productive Mathematics Discussions*. Reston, VA: National Council of Teachers of Mathematics and Corwin Press.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10, 313-340.

Stein, M. K., Grover, B. W., and Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal,* 33, 455–488.

Zembal-Saul, C. (2009). Learning to teach elementary school science as argument. *Science Education,* 93, 687 – 719.

Understanding the Relationships Within and Between Constructs of a Learning Progression: Combining Multidimensional Item Response Modeling and Latent Class Analysis

Jinnie Choi and Ravit Golan Duncan
Rutgers University, 10 Seminary Place, New Brunswick, NJ 08901
Email: jinnie.choi@rutgers.edu, ravit.duncan@gse.rutgers.edu

Abstract: Learning progressions are hypothetical models of student learning in a domain over extended periods of time. In many cases these progressions describe multiple 'big ideas' or constructs. Relationships between these constructs, i.e., how development along one might affect the other, are difficult to ascertain. Such relationships can be described from the perspectives of either item characteristics or student abilities. Existing methods of analyses focus predominantly on the 'item-side' of the equation and much less research addresses construct relationships from the 'student-side'. In this study, we supplemented a Multidimensional Item Response Modeling approach with a Latent Class Analysis to more fully explore both within and between-construct relationships. We analyzed student written responses (n=317) to 31 ordered-multiple-choice items targeted at five constructs in a genetics learning progression. We present our finding with the goal of comparing and contrasting the types of inferences that can be made with both measurement approaches.

Introduction

Learning progressions (LPs) embody a developmental approach to learning by describing productive paths that students might take as they develop progressively more sophisticated ways of reasoning in a science domain over extended periods of time (Alonzo & Gotwals, 2012; Duncan & Hmelo-Silver, 2009a; NRC, 2007). Some LPs map out progress along only one core idea (e.g. Shavelson et al., 2005; Rivet & Kastens, 2012), this is termed a construct map (Wilson, 2005). Alternatively a LP can map out progress along several constructs simultaneously, thus the LP includes multiple construct maps (e.g. Duncan, Rogat & Yarden, 2009b; Jordan & Duncan, 2009; Plummer & Krajcik, 2010). A basic assumption of LPs is that within a construct map students' progress from less sophisticated levels to more sophisticated levels. However, the progress is rarely neat and linear, and diagnosing the level at which a student is reasoning can be challenging. For example, several researchers (Gotwals & Songer, 2010; Steedle & Shavelson, 2009) have pointed to the problem of a 'messy middle', in which students at the middle levels are relatively inconsistent in their reasoning on items of the same relative difficulty. This suggests the likelihood of multiple non-linear paths that students may take to reach the upper level of a LP (Steedle & Shavelson, 2009).

The relationship in progress along multiple construct maps is even more complex and can take many forms. Wilson (2009) offered several representations of how progress along multiple constructs may occur: (a) progress rates along multiple constructs may be very similar such that students progress from one level to the next along multiple constructs at the same time, i.e. the construct maps are aligned; (b) progress along one construct depends on first attaining some level of understanding along another different construct, i.e. the construct maps are staggered; or (c) two or more construct maps may 'feed' into another more sophisticated construct map, i.e. a combination of aligned and staggered maps.

Many researchers (e.g. Brown, Nagashima, Fu, Timms, & Wilson, 2010; Hadenfeldt, Neumann, & Liu, 2013; Anderson, Gotswals & Songer, 2010) have analyzed the relationships among the levels of performances within and between constructs using the multidimensional item response modeling (MIRM) approach (Adams, Wilson, & Wang, 1997; Wilson, 2013). This approach juxtaposes student abilities and individual item difficulties on the same logit scale. In the MIRM approach, inference about the validity of the proposed levels in a construct map depends mostly on how items behave given student abilities. On the 'item-side' one can calculate, for each level, the threshold point for which students have a 50% probability of achieving that level of understanding or higher (termed Thurstonian thresholds) (Wu & Adams, 2007). These thresholds are useful for inferring relative difficulties of moving from one level to the next within a construct map (Wilson & Draney, 2002). Thurstonian thresholds can also be used to infer about relative difficulties of specific levels across constructs. For example, the level one thresholds for items measuring construct X may be similar to, lower, or higher than the level one thresholds for items measuring construct Y.

On the other hand, inferences about the levels of LPs, in particular within constructs, are less informed by the 'student-side' results. The MIRMs provide student ability estimates that are normally distributed. For a five-construct test, each student gets five estimated abilities on each of the constructs and the distribution of these ability estimates for the five constructs may have different means and variances. Thus, within and between construct comparisons can be made, but only in a general distribution sense. That is, whether students can be

classified into the levels of LPs is a question that MIRM approach does not directly answer. The abilities of the students are modeled and estimated to be on a continuous scale, it therefore becomes problematic to later classify students onto discrete levels of performances on LPs. This is because MIRM is based on assumption that the students are in a homogeneous group that shares a particular performance pattern on sets of assessment items; the approach assumes that all students at a particular level reason in the same, consistent, manner. Consequently, it becomes difficult to identify and understand the characteristics of the 'messy middle' classes of students. We explore whether the relationships among the levels of LP within and between constructs can be more fully explained when we supplement and bolster the MIRM approach with the missing component: providing student-side information that matches the discrete nature of the LP levels, and enables within and between construct comparisons of level dynamics in LP.

We use latent class analysis (LCA; Lazarsfeld & Henry, 1968) to provide the student-side information for our analysis of the genetic learning progression (Duncan et al., 2009b). LCA examines if the cases (e.g., students) can be placed into multiple latent groups or classes based on their response patterns. Application of LCA is not new in LP research; Steedle and Shavelson (2009) employed LCA to evaluate whether there are groups of students who perform as expected by an LP for force and motion (Alonzo & Steedle, 2009). Their results suggested that students at the lower and upper levels of the progression reasoned relatively systematically across items, however students at the middle levels often did not reason consistently and were difficult to diagnose as reasoning at a particular level. Similarly to Steedle and Shavelson we use LCA to examine whether the assumptions of our LP match the patterns of the identified classes from data. Moreover, we are particularly interested in the dual use of LCA and MIRM to provide a more complete student and item-side perspectives on the expected performance within and across the five constructs of the genetics LP. Our research questions are thus: (a) what inferences can one draw from the MIRM analysis about the relationships between levels within a construct and between levels across constructs? (b) What inferences can one draw from the LCA analysis about the relationships between levels within a construct and between levels across constructs? (c) In what ways are findings from these approaches congruent, conflicted, or enhanced by each other?

Genetics Learning Progression and Assessment Design

The genetics learning progression is organized around two core questions in the domain: (a) how do genes influence how we, and other organisms, look and function? And (b) why do we vary in how we, and other organisms, look and function? There are eight big ideas associated with these questions. In our current work we are focusing on five of them: (1) Construct A: all living things have genetic information that is organized hierarchically; (2) Construct B: the genetic information specifies proteins structure; (3) Construct C: Proteins have a central role in the biological function of living things and are the mechanism that connects genes and traits; (4) Construct E: Organisms reproduce by transferring their genetic information to the next generation; and (5) Construct F: There are patterns of correlation between genes and traits, and there are certain probabilities with which these patterns occur. Each construct is mapped out across four levels of growing sophistication. Progress along the progression entails developing more sophisticated understandings of these constructs as well how they relate to each other. A detailed description of the progression can be found in Duncan et al. (2009b).

The genetics LP, as originally described, did not provide any conjectures about how development along one or more constructs might affect development along others, as the research base was insufficient to inform such assertions. In recent work we discussed some tentative dependencies between two of the constructs (B and C) and showed that understandings of these constructs develops mostly independently and in parallel (Shea & Duncan, 2013). In this study we attempt to explore such relationships from multiple perspectives, using a larger sample, and with more powerful measurement models. Towards this end we developed a written assessment comprised of 31 ordered-multiple-choice (OMC) items corresponding to the five constructs and their four levels of understanding. In OMC items different response options are linked to levels of conceptual understanding (Briggs, Alonzo, Schwab & Wilson, 2006; Briggs & Alonzo, 2012); items are scored using partial credit models and thus provide more information about students' level of reasoning than traditional multiple-choice items.

Methods

Data and Instrument

The 31 OMC written assessment was administered, over a two-week period, by six participating teachers in 17 biology classrooms (n=317) at a suburban high school in eastern United States. The school consisted of 47% African American, 22% White, 19% Hispanic, and 11% Asian students; 34% of the students were eligible for free or reduced lunch. Among the 17 classrooms, 7 classrooms were higher-performing classrooms or 'honors' (n=164), and 10 classrooms were regular-level classrooms or 'labs' (n=153). Prior to data collection, the six participating teachers implemented the district's eight-week genetics unit covering typical high school level genetics concepts in classical and molecular genetics.

As described above the 31 OMC items were designed to gauge students' understanding of five constructs in the genetics LP. Most of our items included response options that mapped onto 2-4 levels of a particular construct. Overall, at least 3 items, and most often 5 items, measured each level in each construct. Assessments were administered in two comparable forms with the same set of items differently ordered on each form. Across all constructs, the response options mapped onto Levels 0 through 3 of the genetics LP, as well as Level '-' which refers to distractors unrelated to any specific level on the LP. Table 1 shows the actual distribution of the items across constructs and levels, for items that had valid level-scored data. For example, although in the assessment we had more than 3 items that were designed to measure Level 3 for construct B, from students' actual responses we had Level 3 answers for only one item among the three construct B items.

Table 1. Actual number of items across constructs and levels of the genetics LP

	Level – (irrelevant)	Level 0	Level 1	Level 2	Level 3
Construct A	1	2	7	4	3
Construct B	1	5	4	4	1
Construct C	4	8	10	8	6
Construct E	0	0	5	4	3
Construct F	3	4	12	10	7

Analysis

We performed MIRM and LCA analyses separately and consecutively to answer the research questions. The primary goal of the MIRM was to estimate the difficulties of individual items and the abilities of students on the same scale. We took a confirmatory approach that assumes ordered response categories and correlations between constructs. The model also assumes that each of the items measures one of the five constructs in genetics LP. Within each construct, the responses to items are independent and have a Bernoulli distribution.

We used multidimensional random coefficient multinomial logit (MRCML) model (Adams, Wilson, & Wang, 1997) for polytomous data to estimate model parameters. Three types of results are provided by MIRM analyses. First, on the 'item-side', we estimate the difficulties of individual items. In particular, we calculated Thurstonian thresholds for each level of each item. Note that the number of thresholds is one minus the number of levels that the item is measuring. Next, on the 'student-side', for each individual student we estimated the abilities on five correlated constructs. Also, the correlations and covariances among the constructs are estimated. Finally, we calculated other relevant statistics such as deviance, EAP reliability, separation reliability, etc. MIRM analysis was performed using ACER ConQuest IRT software (Wu, Adams, Wilson & Haldane, 2007). ConQuest uses conditional maximum likelihood algorithm to estimate model parameters. The EM algorithm was terminated at the convergence criteria of 0.01 after 16 iterations.

The primary goal of the LCA was to examine whether the students can be placed into multiple latent groups or classes. Note that in using the latent class models we take an exploratory approach that does not assume ordered classes. That is, the resulting classes do not necessarily match the order of levels of the constructs. However, because the response categories were scored following the ordered levels of each construct, we can take a confirmatory perspective in examining if the response patterns reveal higher or lower level performances in some classes. i.e. do the classes differ in ability. We used latent class models (Lazarsfeld & Henry, 1968) for polytomous data to determine classes of students. The varying number of classes was incorporated in the model assumptions, which can be tested by comparing posterior fit statistics. The model assumes that within each class, the items are independent and have a Bernoulli distribution. Given the distributional assumptions of the items, we can express the likelihood of any set of occurrences.

Three types of results are provided by latent class analysis. First, we estimated the probability πick that the response for each item, answered by students from each of the specified number of classes, is equal to a certain response category. Next, we estimated the posterior probability that each of the students falls into each class. For each student, the sum of these probabilities across classes equals one. Finally, we estimate the posterior probability that each student belongs to each class. Latent class analysis for polytomous outcome variables was performed using poLCA (Linzer & Lewis, 2011), a software package implemented in the R statistical computing environment. poLCA uses the EM algorithm to estimate model parameters. The known problem of the EM algorithm is that a local maximum of the log-likelihood function can be found depending on the initial values. To avoid local maxima problems, we ran poLCA 100 times for each model to ensure the results are based on the model with the global maximum likelihood. We selected the results that occurred more than 65 times out of 100 runs. Since one statistic is never a perfect measure of model fit, we looked at three statistics to assess the model fits of the global solution. The first- Log likelihood is a function of the observed responses for each student and the model parameters. The second- AIC is a measure of goodness of fit of a model that considers the number of model parameters; and the third- BIC is a measure that considers not only

the number of model parameters but also the sample size. Preferred models are those that minimize values of the BIC and/or AIC. We also looked at Pearson's chi-squared (X^2) goodness of fit and likelihood ratio chi-squared (G^2) statistics for the observed versus predicted cell counts. Larger values of X^2 and G^2 indicate that the particular model fits the data better.

Results

Examining LP Level Dynamics Within and Between Constructs using MIRM Results

As noted earlier we began with the MIRM 'item-side' analyses to obtain item difficulty estimates and Thurstonian thresholds for item level scores. In Figure 1, the Wright Map depicts the core advantage of MIRM approach- inference about student performance, item difficulties, and levels of LPs can all be made on the same logit scale. On the left five panels, the estimated distributions of student abilities for the five constructs are shown as bell curves. On the right five panels, the estimated thresholds of each level score of each item is shown with colored dots: level 1 thresholds with red, level 2 with green, and level 3 with blue. The gray columns indicate items, and colored horizontal lines indicate average thresholds for levels within constructs.

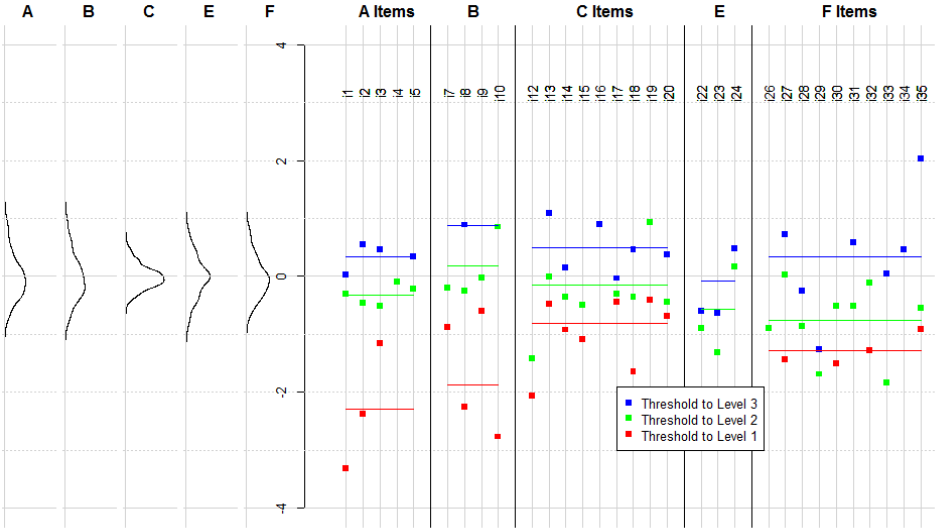


Figure 1. The Wright Map for genetics learning progression.

There are several interesting inferences that can be made from this Wright map. First, there are differences in the level thresholds across constructs. The lowest level 1 threshold across all constructs was for construct A. This lower threshold value on the logit scale implies that it was easier for students to show understanding at level 1 or above on construct A, compared to other constructs. Second, some constructs are overall easier than others. For example, construct E has the lowest threshold for both levels 2 and 3, and appears to be the easiest construct; whereas construct B appears to be the most difficult to master. Third, there are differences in the spread of level thresholds across constructs. For example, constructs A and B have a larger spread compared to C and F, suggesting that there is a larger difference between the understandings at each level of construct B. The demarcation between levels is greater for construct B compared to most of the others. Comparing level threshold spreads within a construct suggests that some 'jumps' from one level to the next are harder than others. In construct A, the difference between level 1 and levels 2-3 was the greatest. Thus while attaining a level 1 understanding on construct A is relatively easy, moving up to a level 2 understanding is much harder compared to similar moves for the other construct.

In reviewing the other side of the Wright map, the 'student-side' results of MIRM afford less inference about the levels of learning progression within or across constructs. While we can compare estimated student ability distribution between constructs, these are not particularly nuanced. We can see, for example, that students performed more similarly to each other on construct C (tighter curve) than on other constructs. However, we cannot compare the location of the distributions between constructs because the mean of the five distributions were all fixed to zero in order to allow the mean of the item difficulties to vary in the MIRM estimation. Thus, the 'between' construct comparison relates to the overall pattern, and is not about relationships between the levels across constructs. Unless standard setting and accompanying student level diagnosis are performed ad hoc (Wilson & Draney, 2002), there seem to be fewer inferences to make about level dynamics using student results from the MIRM analyses. We next present our findings from the LCA approach regarding the relationships of levels within and across constructs.

Examining LP Level Dynamics Within and Between Constructs using LCA Results

LCA allows one to identify classes of students who reason similarly across the entire assessment or individual constructs. To identify how many different classes of students exist, we fit multiple models, from two to five classes. Our results suggest that for constructs A, B, C and E, the model with two classes fitted slightly better than models with more classes. For construct F, the three-class model fitted better than two- or four or five -class models. Given that most constructs have four levels, finding the best fit in a 2-class model was unexpected. One potential explanation is that the honors and lab students function as two distinct classes that overshadow other more subtle distinctions. We subsequently decided to take the four-class solutions for all constructs and look at the characteristics of the classes in detail.

In the four-class model we estimated the conditional probability that a student in a class responds to an item with a certain response category (i.e. at a particular level). For each student and each item, the probabilities of responding to all categories are assumed to sum to one. Figure 2 summarizes these conditional probabilities. In the left graph of Figure 2, we show the results from the four-class solution with construct A items. Each of the four panels on the graph represents a predicted class. Note that class number does not necessarily match the order of the levels in LP (i.e. class 1 is not necessarily students reasoning at a level 1 on this construct). On the X axis we have five A items, and on the Y axis, we have the conditional probability of getting a certain level score for an item. Each bar represents an item and each color represents a level score for the item. The height of each segment in the bar represents the likelihood of students in that class obtaining a particular level score for the specific item. Ideally, we expect to see is that classes are different in terms of the proportion of different colors they have (i.e., there is one dominant level for each class across the items); then one can argue that the classes reflect the characteristics of the ordered levels in LP. Here, the figure shows that Level 3 (purple) responses are dominant for class 1, Level 2 (blue) is dominant for class 3, Level 1 (green) is dominant for class 4 and Level 0 (red) is dominant for class 2. Based on actual summation of the probabilities, the most dominant level in class 1 is Level 3, for class 2 it is Level 0, for class 3 it is Level 2 and for class 4 it is Level 1. This is a relatively 'clean' match between levels and classes. Predicted class memberships, estimated by the modal posterior probability, show that 26.5% of students belong to class 1, 16.1% belong to class 2, 11.7% to class 3, and 45.7% of the students belong to class 4.

In comparison with the cleaner LCA analysis for construct A, the right graph of Figure 2 presents the 4-class analysis for construct F, which seems to be the messiest among the five constructs. Visually, Level 3 (purple) is dominant for class 2, Level 2 (blue) is dominant for class 2,3, Level 1 (green) is dominant for class 4, Level 0 (red) is dominant for class 1,3,4, and the unrelated Level '-' (orange) is dominant for class 1. Level '-' represents item response options that were simply distractors and did not map onto any specific level on the LP. The summation of probabilities reveals that most dominant level for class 1 is Level 2, for class 2 is also Level 2, for class 3 is again Level 2, and for class 4 is Level 1. Predicted class memberships show that 43.5% of students belong to class 1, 48.0% belong to class 2, 5.1% to class 3, and 3.5% of the students belong to class 4.

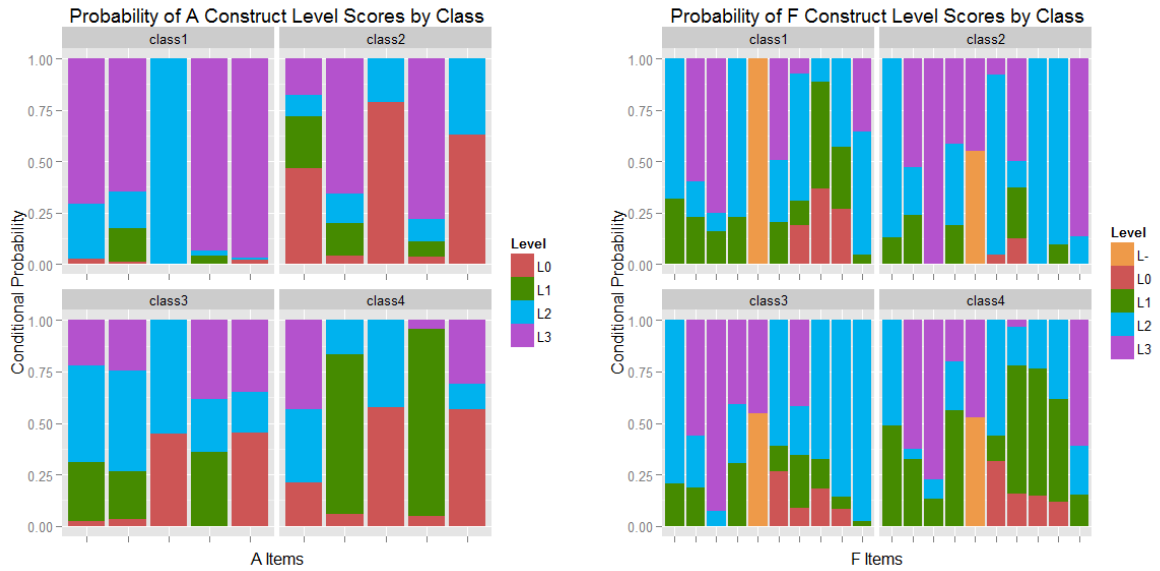


Figure 2. Estimated conditional probability of constructs A and F level scores by four latent classes.

Table 2 shows the classification of the student abilities across all five constructs, given the four-class solutions. Across the five constructs, how student abilities are classified into the four levels of the LP was clearly different. The best classification result was observed for construct A: each of the four classes in construct A matched well, also proportionally, with each of the four levels of the construct (shown by the

dominant level and its proportion among all answers.) A majority of students were in class 2 (43.0%), of which student responses to the construct A items were mostly at level 2 (39.4%). However, the classification results did not match well with the four levels for all other constructs. For construct B, the four classes were characterized by only two levels of the construct: level 1 and 2. A majority of students were classified into class 3 and class 4 (48.9% + 28.4% = 77.3%), of which student responses to the construct B items were mostly at level 2 (44.9% and 54.9%). For constructs C, the four classes of students were characterized only by two dominant levels: levels 1 (class 4, 7.8%) and 2 (all three other classes, 92.2%). For construct E, the four classes were characterized by levels 1 and 3, with the majority of students at level 3 (64.4% + 21.7% = 86.1%). For constructs F, the four classes were also characterized only by levels 1 (class 4, 34.6%) and 2 (all three other classes, 65.4%).

Table 2. Classification of the students across all five constructs

Construct	Class	Predicted Class Membership	Dominant Level	Proportion of Dominant Level Answers, Across Items
A	1	0.160	Level 3	0.651
	2	0.430	Level 2	0.394
	4	0.217	Level 1	0.337
	3	0.193	Level 0	0.392
B	3	0.489	Level 2	0.449
	4	0.284	Level 2	0.549
	1	0.157	Level 1	0.628
	2	0.070	Level 1	0.383
C	1	0.312	Level 2	0.404
	2	0.525	Level 2	0.354
	3	0.085	Level 2	0.205
	4	0.078	Level 1	0.258
E	2	0.644	Level 3	0.747
	3	0.217	Level 3	0.450
	1	0.095	Level 1	0.601
	4	0.044	Level 1	0.750
F	1	0.111	Level 2	0.377
	2	0.236	Level 2	0.453
	3	0.308	Level 2	0.476
	4	0.346	Level 1	0.352

Overall, LCA, like MIRM, affords making some interesting inferences. First, as noted earlier, the best fitting model is a two-class model rather than a four-class model, suggesting that the proposed levels of the progression may not map neatly, or at all, onto students' actual performance. Second, when using a four-class model we find that some constructs are much messier than others. By this we mean that the classes in some constructs map poorly onto levels (construct F) compared with classes in other constructs (construct A). Often there are clearer class-level association for the highest and lowest performing classes (cleaner mapping onto the least and most sophisticated levels of the construct map) and a much messier middle, a phenomenon that has been previously documented (Gotwals & Songer, 2010; Steedle & Shavelson, 2009). Students in this messy middle tend to reason inconsistently, performing well on some items and less well on others. Third, in some cases one can make comparisons between classes across different constructs. For example, our analysis suggests that the students who are classified in class 3 on construct A are mostly classified in class 1 on construct B (not shown). Both these classes (class 3 in construct A and class 1 on construct B) reason at a level 3 on both constructs respectively. However, making such comparisons between constructs A and F is problematic due to the rather fuzzy distinctions between classes on construct F. Thus the messier the constructs the more difficult it is to compare them and make inferences about cross-construct relationships.

Discussion

Overall, our results suggest that, not surprisingly, the MIRM and LCA analyses together provide more detailed and nuanced information than each alone. We have shown that MIRM provides useful information about how the levels of the items within constructs and across constructs are perceived by students. That is, which items are easy and which are hard, which constructs are overall easier and which are harder. However, MIRM does not

provide much useful information about how certain groups of students within our sample behaved differently, within and across constructs in attaining different levels on the items and consequently constructs. The student ability estimates are provided with an assumption of a continuous scale, not with distinct classes, groups, or levels. We can later classify the students onto levels of LPs, but it depends on 'item-side' threshold results that do not account for difference among student groups. Consequently, MIRM may not be sufficient in understanding the problem of interest: the relationships among the levels of LPs within and between constructs. Our findings suggest that LCA is useful in providing additional, 'student-side', information about the 'messy middle' levels or classes in certain constructs of LPs. However, LCA is less amenable to ranking students' performances or to assess correlation between the performances on multiple constructs. This is because ranking and correlations require continuous data, yet LCA allows classification of students onto a few distinct levels of performances.

The benefit of using multiple measurement models and approaches to studying LPs has been noted by other researchers (Briggs & Alonzo, 2012) and there are several different approaches that have been used besides the more frequently-used MIRM (e.g., attribute hierarchy method (AHM; Briggs, Alonzo, Schwab, & Wilson, 2006), Bayesian networks (West et al., 2010). In this research, we chose to bolster the popular use of MIRM in LP research with the use of LCA in order to more fully explain the relationships among the levels of LP within and between constructs. This was possible because LCA provided student-side information that matches the discrete nature of the LP levels. While LCA, AHM, and Bayesian networks can classify students into discrete classes, LCA is a less diagnostic but simpler approach. With the cost of more detailed diagnostic information, LCA does not require a-priori specification of a matrix that formally associates items and attributes, as in AHM, nor a multitude of conditional probability tables, as in Bayesian networks, Navigating between multiple methodological frameworks for empirical validation of LPs is already a problem for researchers when resources do not allow clear guidance in the pool of methodologies. While a more formal, comprehensive comparison of methods should follow to further inform researchers, this study contributes to the ongoing scholarship on LPs by providing some reasons and guidance for choosing between the MIRM and LCA approaches given a specific research goal. By doing so, this study motivates further discussions about what types of evidence each of these methodologies provides, or not , in relation to different research questions.

The work also highlights some important implications regarding the genetics progression and learning genetics more specifically. For example, there are certain ideas that are easier for students to master than others. In this case we found that reasoning about the hierarchical organization of the genetic information (connection between DNA, chromosome, genes, nucleotides) and the universal nature of the genetic information (all organisms have genetic information that is used by their cells in essentially the same way) was an idea (construct A) that was relatively easy for students to master. However, understanding what the information is about and the cell uses the information, was the hardest idea (construct B). This may be a reflection of the common instructional focus on structure and process (structure of DNA, processes of transcription and translation) rather than on the big idea that genes are instructions for proteins and that proteins are the physical mechanism that generates our traits (Duncan & Reiser, 2007). On the other hand, understanding that parents give half the genetic information to their offspring was an easier idea for students to master and the movement up the levels of this construct (E) involved much smaller conceptual jumps .

There are also interesting differences in the spread of levels of understanding for different constructs. For example, construct C (role of proteins) shows much less variation in understanding across students. Most students are at a level 2. In comparison constructs A and B show greater variation. We believe this reflects differences in the nature of the constructs and the extent to which students have substantive prior knowledge about those ideas. It seems that students may develop understandings of construct C as a result of instruction (recall the data shows were collected after genetics was taught) and thus have fairly similar "party line" understandings of proteins. However, their understandings of the nature of the genetic information (construct B) are likely informed more extensively by prior knowledge from various sources and students may exhibit such understandings on the assessment resulting in a larger spread of student ability for that construct.

In terms of using our analyses can be used to revise the progression the picture is rather fuzzy. The relatively small sample of the study and the fact that the instruction was not based on the expectations of the progression makes drawing clear-cut conclusions difficult. The point about instruction is rather critical. If we assume that students will progress along a hypothetical progression when they experience instruction that supports such progression (i.e. instruction that capitalizes on the developmental constraints and affordances embodied in the progression), then the nature of the instruction experienced becomes a critical part of the equations. If instruction is not designed based on the progression it is not clear whether the expectation for anticipated student progress should hold.

References

Adams, R. J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*, 1-23.

Alonzo, A. C, & Gotwals, A (Eds.) (2012). *Learning progressions in science: Current challenges and future directions.* Rotterdam: Sense Publishers.

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education, 93*(3), 389–421.

Briggs, D. C., & Alonzo, A. C. (2012). The psychometric modeling of ordered multiple-choice item responses for diagnostic assessment with a learning progression. *Learning Progressions in Science*, 293-316.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. R. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment, 11*(1), 33-63.

Brown, N. J., Nagashima, S. O., Fu, A., Timms, M., & Wilson, M. R. (2010). A framework for analyzing scientific reasoning in assessments. *Educational Assessment*, *15*(3-4), 142-174.

Duncan, R. G., & Hmelo-Silver, C. (2009a). Learning progressions: Aligning curriculum, instruction and assessment. *Journal of Research inScience Teaching, 46*(6), 606–609.

Duncan, R. G., Rogat, A., & Yarden, A. (2009b). A learning progression for deepening students' understanding of modern genetics across the 5th-12th grades. *Journal of Research in Science Teaching, 46*(6), 644-674.

Hadenfeldt, J. C., Neumann, K., & Liu, X. (2013). Validating a model of students' progress in understanding the concept of matter. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Rio Grande, Puerto Rico.

Gotwals, A. W., & Songer, N. B. (2010), Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education, 94*(2) 259–281

Jeon, M., & Rabe-Hesketh, S. (2012). Profile-likelihood approach for estimating generalized linear mixed models with factor structures. *Journal of Educational and Behavioral Statistics, 37*(4), 518-542.

Jordan, R. & Duncan, R. G. (2009). Student teachers' images of science in ecology and genetics. *Journal of Biological Education, 43*(2), 62-69.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Houghton, Mifflin.

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software, 42*(10), 1-29.

Mohan, L., Chen, J. and Anderson, C. W. (2009), Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46*(6), 675–698.

National Research Council [NRC]. (2007). *Taking Science to School: Learning and Teaching Science in Grade K-8.* Washington DC: The National Academy Press.

Plummer, J.D. and Krajcik, J.S. (2010). *Building a Learning Progression for Celestial Motion: Elementary Levels from an Earth-Based Perspective. Journal of Research in Science Teaching*, 47(7): 768-787

Rivet, A., & Kastens, K. (2012). Developing a construct-based assessment to examine students' analogical reasoning around physical models in earth science. *Journal of Research in Science Teaching, 49*(6), 713-743.

Shavelson, R., Stanford Educational Assessment Laboratory (SEAL) and Curriculum Research & Development Group (CRDG). (2005). *Embedding assessments in the FAST curriculum: The romance between curriculum and assessment.* Final Report.

Shavelson, R. J. (2009). Reflections on learning progressions. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Shea, N. A., & Duncan, R. G. (2013). From theory to data: The process of refining learning progressions. *Journal of the Learning Sciences, 22*(1), 7-32.

Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching, 46*(6), 699-715.

West, P., Rutstein, D. W., Mislevy, R. J., Liu, J., Levy, R., DiCerbo, K. E., ... & Behrens, J. (2009). A Bayes net approach to modeling learning progressions and task performances. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA

Wilson, M. (2013). Constructing measures: An item response modeling approach. Routledge Academic.

Wilson, M., & Draney, K. (2002). A technique for setting standards and maintaining them over time. *Measurement and multivariate analysis*, 325-332.

Wu, M. L., & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne: Educational Measurement Solutions.

Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, (2007). ACER ConQuest 2.0 [computer program]. Melbourne: ACER.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No.1053953. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Students' Use of Evidence and Epistemic Criteria in Model Generation and Model Evaluation

Ravit Golan Duncan, Carol Tate, and Clark A. Chinn
Rutgers University,10 Seminary Place, New Brunswick, NJ 08901
ravit.duncan@gse.rutgers.edu, carol.tate@gse.rutgers.edu, clark.chinn@gse.rutgers.edu

Abstract: The Next Generation Science Standards and the Framework for Science Education emphasize the importance of engaging learners with the core scientific inquiry practices of modeling and argumentation. Students are also expected to understand the epistemic grounds and norms that accompany these practices. We report on a study in which we engaged middle school teachers, and their students, in model-based inquiry, with particular emphasis on developing models and evaluating competing models using evidence. Analysis of students' written arguments, in the context of an assessment task in genetics, suggests that students use both secondary epistemic criteria, relating to communicative features of models (labels, drawings), as well as primary epistemic criteria, relating to evidence-model fit. Most students used at least one, and often several, provided pieces of evidence to support their arguments. We also discuss some instructional implications and tradeoffs in selecting evidence for such model generation and model evaluation tasks.

Introduction

The recently released Next Generation Science Standards (NGSS) (Achieve, 2013) adopt a view of scientific inquiry as a knowledge-building enterprise that employs a systematic, and evidence-based, approach to building models that explain the world around us (Giere, 2004; Godfrey-Smith, 2006). Scientific models are abstract, simplified representations of important aspects of the scientific phenomenon under study (Doerr & Lesh, 2003). These models are developed iteratively through a process of testing and revision. Evidence, and reasoning about evidence, is at the core of these processes (Longino, 2002). Scientists developing these models operate within a community of research, with continually negotiated norms regarding what counts as good evidence, arguments, and models (Kitchner, 1993; Kuhn, 1977; Latour, 1987).

Over the past two decades, there has been substantial research investigating the learning and teaching of scientific modeling practices (e.g. Driver, Leach, Millar, & Scott, 1996; Grosslight, Unger, Jay, & Smith, 1991; Lehrer & Schauble, 2000; Schwarz & White, 2005; Treagust, Chittleborough, & Mamiala, 2002). However, we still know relatively little about the ways in which students understand what counts as a good model, evaluate the quality of evidence, and relate evidence to one or more explanatory models (Lehrer & Schauble, 2000; Pluta, Chinn & Duncan, 2011; Schwarz *et al*, 2009). For example, Schwarz and White (2005) had students evaluate models using four criteria that were provided to them: accuracy, plausible mechanism, consistency, and utility of models. They found that students who used the four criteria showed a better understanding of the nature of modeling, scientific inquiry, and the targeted physics content compared to students who completed the same instructional unit, but without explicit use of criteria. In our own work we have shown that middle school students are also capable of generating their own criteria for judging model quality (Pluta, et al., 2011). Students' criteria were predominantly about communicative features of models such as models being clear, labeled, organized, and including pictures. However, almost a quarter of the students did note criteria that related to the model's fit with evidence, and almost half claimed that models should explain the phenomenon under study. These findings are congruent with earlier research suggesting that, at least some, high school students see models as important for developing explanations and making predictions, and that they are revised in light of evidence (Grosslight et al., 1991; Treagus et al., 2002). Helping students develop more sophisticated epistemic understandings of models and their role in science is important if we want them to fully grasp and engage with the modeling practices advocated by the NGSS and the Framework for Science Education (National Research Council, 2011). In accordance with the ICLS 2014 theme, developing such epistemologies is part and parcel of learning science and becoming scientifically literate.

In our current project, we have developed a set of epistemic scaffolds to support student engagement with models, evidence and the relationship between them (Chinn & Buckland, 2012; Rhinehart, Chinn & Duncan, in press). There are three core scaffolds that we have used with middle school teachers and students: (a) student generated lists of criteria for model-goodness; (b) Model-Evidence-Link diagrams (MELs), in which students use five different types of arrows (support, contradict, strongly support, strongly contradict, and irrelevant) to connect each piece of evidence to multiple competing models; and (c) evidence rating boxes within the MEL diagrams, in which students record their judgments of the quality of each piece of evidence on a scale from 0 (very poor evidence) to 2 (high quality evidence).

Here we report findings from a modeling-and-argumentation assessment task in which students had to

 © ISLS

first develop their own models of a hypothetical genetic disorder given a few pieces of evidence. They then used a MEL diagram to evaluate, and choose between, two competing explanatory models of that disorder in light of additional provided evidence. Students completed this assessment at the end of a 5-week unit on genetics. Our research questions thus align with our two-fold goal for the task: (a) evaluating the extent to which students can generate mechanistic models of the cellular and molecular mechanisms that underlie genetic phenomena, and (b) evaluating students arguments and, in particular, their use of evidence and epistemic criteria.

Research suggests that students struggle to provide mechanistic accounts of genetic phenomena that explain how our genetic information brings about physical traits (Lewis & Kattmann, 2004; Marbach-Ad & Stavy, 2000). This is, in part, due to the many unfamiliar cellular and molecular entities involved, such as DNA and proteins (e.g. Lewis & Wood-Robinson 2000; Marbach-Ad, 2001;Venville & Treagust, 1998), and in part to the current instructional methods that tend to blackbox the protein-based mechanisms that link genes to traits (Duncan & Reiser, 2007). Several researchers have proposed instructional frameworks and scaffolds to support students in developing mechanistic explanations of genetic phenomena (Duncan & Reiser, 2007; van Mil, Boerwinkel, & Waarlo, 2011). Implementations of such scaffolded curricula have met with some success at the middle and high school level (Duncan & Tseng, 2011; Duncan, Freidenreich, Chinn & Bausch, 2011). Introducing these ideas early on is the key to supporting more robust understandings by the end of schooling. This view is reflected in the NGSS, which, unlike prior iterations of the standards (NRC, 1996), introduce the relationship between genes, proteins, and traits at the middle school level (MS-LS3-1).

Therefore, as part of our study of modeling and argumentation with middle school students we have developed a five-week unit in genetics that focuses predominantly on Mendelian genetics but also addresses the link between genes, proteins and traits in the context of genetic resistance to HIV (described in detail below).

Methods

Study Context

The study was conducted in a relatively large suburban 6th and 7th grade middle school (approximately 1450 students) in the Northeast. The majority of the students in the school were Caucasian (61%) with a large minority of Asian students (28%), and small minorities of Hispanic (6%) and African-American (5%) students. Approximately 14% of the students were eligible for free and reduced lunch. Four 7th grade teachers participated in the study with their approximately 400 students. Participating teachers implemented five months of instruction using materials we developed jointly, interspersed with their own materials. The study involved two conditions: (a) the treatment condition included a consistent and explicit focus on developing and using criteria for model-goodness that were student-generated and revised periodically throughout the duration of the implementation, and (b) the control condition in which there was no explicit and public focus on model-goodness criteria. Both conditions used the MEL diagrams and the evidence rating boxes.

The implementation study began with a set of activities designed to: (a) introduce students to the norms of argumentation discourse (giving reasons, disagreeing nicely, etc.); (b) engage students in the generation of a consensus list of model goodness criteria (*only* in the treatment condition); and (c) introduce students to the MEL diagrams and evidence rating boxes and procedures. These introductory activities were followed by a unit on cell organelles in which the teachers used study materials for two of the organelles they taught- chloroplast and nucleus (they used their own materials for other organelles typically covered in this unit). Following the cell unit was the 5-week genetics unit. This unit began with several lessons about Mendelian genetics during which students developed model for the "rules" governing inheritance patterns they observed in pedigrees, and then learned the relevant terminology and algorithms (Punnett squares) used to describe inheritance patterns and the probabilities of particular gene and trait combinations. The unit then turned to molecular genetics and the remaining lessons dealt with inherited resistance to HIV (some people are not susceptible to HIV infection because the virus cannot enter and infect their white blood cells). This part of the unit also included a set of teacher-planned activities about the structure and function of DNA. The HIV lessons involved the evaluation of two competing models for the genetic basis of HIV resistance that linked a mutation in a gene to the resistance trait. In one model the mutated gene gives instructions for making a novel protein that attacks the virus and thus confers resistance; in the second model the mutated gene results in a missing membrane protein that is normally used by the HIV as an anchor (necessary for infection of the cell by the virus). The second model approximates the currently acceptable mechanism for HIV resistance. As in most of our instructional activities, students evaluated these models against multiple pieces of evidence, such as a simplified summary of a scientific study about the presence or absence of particular proteins in the cell membrane of normal and resistant individuals. Students wrote extensive arguments in support of their chosen model; students in the treatment condition were encouraged to refer to the model-goodness criteria as they developed their arguments.

In this paper we report about one of the teachers and her 90 students: 40 in two class sections assigned to the treatment condition, and 50 in three class sections in the control condition. The teacher was untenured and in her second year of teaching. She held progressive views of teaching and was eager to engage her students

with model-based inquiry instruction. She enacted the genetics unit with high fidelity based on our field notes and video tapings of her lessons.

Data Sources and Analysis

The written assessment described herein was given to all students at the end of the genetics unit. The assessment was comprised of two tasks both involving a scenario of a hypothetical skin disease "DEB" in which individuals have blisters in their skin. In the first task, students were asked to explain using pictures and words, "…what you think is happening inside the bodies of people with DEB?" They were also provided with three pieces of evidence related to DEB: Evidence 1 described the inheritance pattern of DEB, Evidence 2 compared samples of healthy skin with skin from DEB patients, and Evidence 3 provided a diagram of healthy skin showing the layered structure. We analyzed students' models to ascertain whether they included an explanation that linked genes to the trait (blisters) via a protein-based mechanism, and whether they used the provided evidence in their models. Students were then asked to critique their model: "How good do you think your explanation is? Give at least four reasons for your answer." Responses to the self-assessment prompt were examined in terms of how students rated the quality of their models, the kinds of model-goodness criteria they referred to in their evaluation of the model, and whether they identified any shortcomings of their models.

In the second task, students were presented with two explanatory models of DEB that provided a mechanism linking a gene to the blistering (see Figure 1).

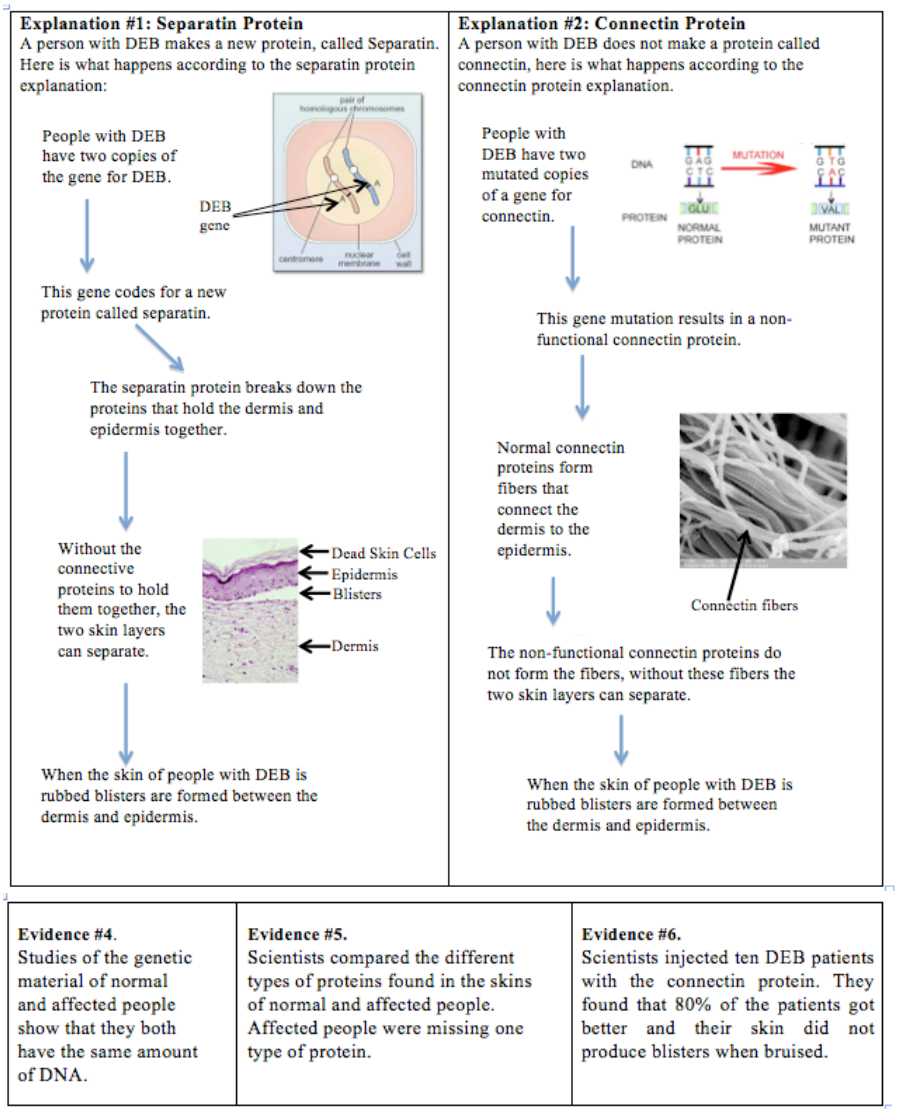


Figure 1. Model-Evaluation Task

The first model (the separatin model) postulated a mutated gene coding for a novel "separatin" protein that caused the skin layers to separate resulting in blisters. In the second model (connectin model) the mutated

gene results in the lack of a "connectin" protein that normally holds the skin layers together. The second model, lack of a protein, is the correct explanation of the real disorder on which the DEB scenario is based. Students initially chose which model they believed to be correct, and then read three additional pieces of evidence shown in Figure 1: Evidence 4 stated that both normal and affected individuals have the same amount of DNA (this evidence supports both models and is essentially irrelevant to choosing between them); Evidence 5 indicated that scientists discovered that affected individuals are missing one type of protein (this evidence was intended to support the connectin model); and Evidence 6 described a study in which scientists injected 10 affected individuals with connectin and 80% of the patients got better (this evidence supports the connectin model but has a rather small sample size). Students were then prompted to reconsider their choice of models in light of the evidence and write a reasoned argument to support their choice.

We analyze these arguments using coding schemes adapted from prior research (Dianovsky, Duncan & Chinn, 2013) to capture the quality of student arguments in terms of students' use of evidence including: (a) how did they interpret the evidence, (b) how many pieces did they cite, (c) did they explain the link between the evidence and the model, (d) did they address counterevidence (if they chose the incorrect separatin model), and (e) did they include any counterarguments against the competing model.

Results and Discussion

Students' Models of DEB

The student-generated models of DEB were mostly phenomenological and did not include any protein-based explanations of DEB (see Figure 2a). Only one of the 90 participating students provided a model that included a postulated mutation in a gene resulting in missing protein that would normally connect the dermis and epidermis skin layers preventing blistering (see Figure 2b). These results are fairly disappointing and we were surprised that none of the other students provided a mechanistic explanation. It may be that they did not generalize the role of genetic mutation and proteins in genetic phenomena from the HIV example taught in the unit. Our prior research did suggest the need for multiple examples and support in generalizing the gene-protein-trait schema (Duncan, et al., 2011), however, due to time constraints with the genetic unit in this study, we did not develop these additional activities. It seems that despite an emphasis on genes being instructions for proteins in the nucleus lesson (taught before the genetics unit), and a similar emphasis in the teacher-generated activities about DNA (central dogma), students did not develop a generalized schema that they could apply in other contexts. It is also the case that the evidence we presented in this first task (Evidence 1-3) did not deal with proteins and thus students were not compelled by the evidence to introduce proteins into their models.

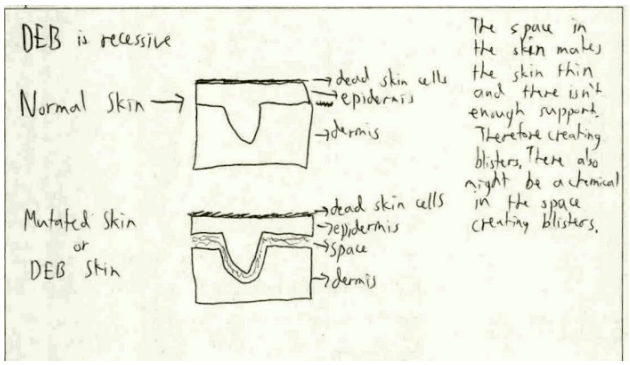


Figure 2a. Typical Student-Generated Model of DEB

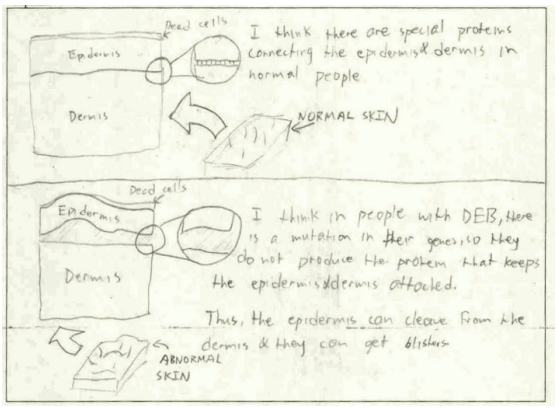


Figure 2b. Mechanistic Model of DEB

Given that students' models were not truly explanatory, we were interested in seeing how they evaluated their own models. We found that most students rated their models positively (noting that the model is very good, good, or OK). These students often cited several of their class model-goodness criteria in justification of their response, including clarity, having drawings, labels, and fit-with-evidence. Table 1 illustrates the criteria and frequency of citation by students. Note that the "I used the evidence" criterion was fairly prominent, and given that the evidence was at the phenomenon level, these students are sensibly citing the criterion. Many students essentially re-drew the evidence showing the two skin layers and the blisters. Thus for all intents and purposes they did use the evidence. Interestingly, there seem to be no significant differences between the treatment and control conditions. Both groups cite the same criteria in support of their models, even though the control condition did not develop public criteria for model-goodness.

Table 1: Justifications given for students' self-assessment of model.

Criteria Referenced by Students in Their Evaluation	Number of Responses (%)	
	Control	Treatment
I explained my answer	14 (28%)	15 (38%)
I used the evidence	15 (30%)	16 (40%)
I drew a picture/diagram	26 (52%)	24 (60%)
I used detail	4 (8%)	1 (3%)
I used labels	7 (14%)	9 (23%)

While none of the students cited their class criteria about mechanism as a commendable aspect of the model (relevant criteria from their class lists were: "shows logical process", "shows sequential steps", "has mechanism"), several students did cite a lack of mechanism as a problem with their model. There were 13 students who did not evaluate their model positively. Of these, eight argued that their model was not great because, " I did not explain my reasoning on what I think is happening inside the bodies of people with DEB", "I am very unsure with how the blisters are formed", "I did not say how it creates blisters." Hence, for these students, the criteria lists helped to highlight a critical gap or shortcoming of the model. The students were unable to address the gap but did acknowledge it, an important epistemic achievement in itself. That the remaining 79 students did not identify this shortcoming is troubling. Then again, since most of the class criteria focused on communicative features of the models it is not entirely surprising that students concluded that their models were adequate if they clearly portrayed the phenomenon and addressed the evidence.

Students' Arguments for the Best Model

We next analyzed students' written arguments in support of their chosen models (separatin or connectin explanations). Overall, 63% of the students chose the correct connectin model, 32% chose the separatin model, and the remaining 5% were undecided. In their arguments, 71% of the students cited at least one piece of evidence in support of their model choice. With the majority of those (49 students) citing two or three pieces of evidence. Twelve of the 90 students explicitly cited the quantity of evidence supporting the connectin model as the reason they chose it. Students' use of evidence varied from merely noting a single piece of evidence, to discussing several pieces of evidence and explaining how the evidence supports the model (justifying evidence-model link). Table 2 illustrates students' use of evidence in their arguments. Note that categories are not mutually exclusive and students' arguments could be double coded. Overall, there seem to be no significant differences between the control and treatment condition, with one exception: twice as many students in the control provided explanations of how the evidence supports the model. We are not sure why this is the case. One possible explanation is the focus on fit-with-evidence and justification was, for some reason, made more salient in the control classes. We currently do not have evidence to support or refute this conjecture.

Table 2: Students Use of Evidence

Nature of Evidence Citation	Number of Responses (%)	
	Control	Treatment
Student does not discuss any piece of evidence	15 (30%)	12 (30%)
Student cites at least one piece of evidence	36 (72%)	28 (70%)
Student explains how the evidence supports the model	21 (42%)	10 (25%)
Student discusses how evidence relates to the competing model	9 (18%)	9 (23%)
Student mentions the quality and/or relevance of the evidence	3 (6%)	4 (10%)
Student notes amount of evidence as contributing to the choice	5 (10%)	7 (18%)

Across both conditions, there were seven students who noted more than one piece of evidence, provided a justification linking the evidence pieces to the model, and considered the competing model in their argument — a fully articulate argument and rebuttal:

> I think the connectin explanation is better. In Evidence 2 it says that when scientists studied normal and affected skin samples. The affected skin had large gaps between the dermis and epidermis just like explanation 1 says. In evidence 5 it says that they compared the proteins and noticed a protein was missing. In this explanation it's saying the connectin protein is missing. In evidence 6 people with DEB are injected with the connectin protein and their skin

> became better. If they had separatin protein it would just break the protein again. This shows that they just never created the protein. That is why the connectin explanation is better.

This student mentions three pieces of evidence (2, 5 and 6) that all support the chosen, and correct, model. The student further explains how the evidence pieces support the model. These evidence pieces point to a protein being missing, which is the core difference between the two models. The student understands this key distinction and notes that evidence 6 thus contradicts the separatin model ("If they had separatin protein it would just break the protein again"). This is a well-articulated, evidence-based, and justified argument and counterargument. Interestingly, five of the 29 students who chose the incorrect model (separatin) were also able to provide arguments that used evidence to refute the, actually correct, model:

> I think the separatin protein explanation is better: (1) Evidence 5 shows how people affected are missing a protein, like in the model; (2) Evidence 6 agrees with the other model, but is only done with 10 people; (3) it [separatin model] is clearer to me and makes way more logical sense.

This student interprets evidence 5 as supporting the incorrect model and then essentially dismisses the conflicting evidence (6) due to sample size. This student, who was in the treatment condition, was one of two students who discussed evidence quality, such as sample size, in their argument. The misinterpretation of evidence 5 was more common (24 students) and rather interesting.

Evidence 5 stated: "Scientists compared the different types of proteins found in the skins of normal and affected people. Affected people were missing one type of protein". Our intent with this evidence was to support the connectin model, which stated that affected individuals are missing the needed connectin protein. Many students did interpret the evidence accordingly:

> Evidence 5 contradicts explanation #1 because evidence 5 says that they are missing one type of protein but explanation one says that they get a *new* protein, which means that have an extra one, they are not missing one.

However, since the separatin model noted that the new separatin protein breaks down the connective protein that holds the layers together, several students interpreted the evidence in accordance with that model:

> Evidence 5 states that a protein is missing. This supports explanation #1 [separatin] because of how explanation 1 stresses that the separatin protein breaks down the protein which means it's not there therefore it supports explanation #1 because the protein is broken down.

While the first interpretation is correct, the second interpretation is sensible. In hindsight, it is clear that this evidence was problematic. Yet, this situation also highlights a design challenge: identifying evidence for evidence-based tasks that is neither too straightforward and simple, nor too ambiguous and open to multiple interpretations. On the one hand, scientific evidence can often be interpreted in multiple ways and understanding this point is an important epistemic achievement. Helping students learn that models are often under-determined by evidence can, and should, be an instructional goal that is facilitated by well-designed materials that present students with ambiguous and controversial evidence. On the other hand, there are specific content goals associated with curriculum materials and in these model-evaluation tasks it is necessary to craft the body of evidence to support the scientifically normative model. In this case, the ambiguity allowed a substantial number of students to choose the erroneous model and support it with evidence. We have no simple solution to this tradeoff in design, but it is a relevant and recurring tradeoff that requires careful consideration.

Conclusion and Implications

Returning to our research questions: (a) can students generate mechanistic models of the cellular and molecular mechanisms that underlie genetic phenomena? And (b) in what ways do students use evidence and epistemic criteria in their arguments? Our findings suggest that overall students were not able to generate mechanistic explanations of the cellular and molecular basis of the genetic phenomenon described in the assessment. Their models were mostly at the phenomenological level and essentially reiterated the symptoms. Evidently, the curriculum as designed and enacted was insufficient in helping students develop a more generalized schema of genetic mechanisms that they could apply to novel contexts. This finding underscores a core implication- that it is essential to help students develop generalized models/schemas of mechanisms in the discipline, and that multiple examples are likely needed to support the development of such generalizations.

In terms of the second research question, it seems that most students demonstrated awareness of some core epistemic criteria for good models and arguments. In evaluating their own models, most students referred

to secondary epistemic criteria related to communicative features of model (labels, clarity, pictures, etc.), and about a third mentioned the primary epistemic criteria of fit-with-evidence. This was the case for both study conditions, which is somewhat surprising since there was no explicit focus on criteria in the control condition. However, as we have shown in prior research (Pluta, et al., 2011), students are capable of coming up with both secondary and primary epistemic criteria for good models without much scaffolding. While students in the control condition did not develop public criteria lists, it seems that the constant discussion of evidence and models likely helped them develop a set of implicit criteria that they used in this task. In contrast, students in the treatment condition did not seem to develop substantially more sophisticated criteria. We wish to caution, however, against over generalization of these initial findings. The work reported here is based on assessments from one of four teachers and on a single task. It may be that this task context afforded less opportunity for students to demonstrate their developing epistemic prowess. In a prior study with middle school teachers from a different school, who had more experience with model-based inquiry and modeling criteria, we have shown that a focus on a related set of criteria (criteria for good evidence and criteria for determining evidence-model relationships) did result in significant gains in argumentation. Students' arguments included more explicit justification of how evidence related to the model, discussed quality of the evidence more often and in more detail, and used more evidence, including counterevidence for the competing model (Dianovsky, et al., 2013). Others have also had similar success in focusing student attention on epistemic criteria (Schwarz & White, 2005). We suspect that while the teachers in the study did engage students in the development, revision, and use of criteria, these criteria lists remained rather intuitive and superficial. The implication here is that to reap the benefits of explicit and public engagement with epistemic criteria, it is important to move beyond what students can do on their own and to really deepen, expand, and enhance their initial lists. In particular, progressing beyond communicative criteria to those that deal with more subtle aspects of model-evidence-fit, explanatory nature of models, accuracy, etc. appears to be essential.

Our findings also suggest that students were able to use evidence, provide reasons and justify their claims, at least to some extent. Almost half of the students in our study used multiple pieces of evidence to support their claims, a third provided justifications that explain how the evidence related to the model, and almost a quarter provided counterarguments against the competing model. These findings echo and extend research of others who have shown that, with proper support, students can use and internalize argument structure, and develop better evidence-based arguments (Duschl, 2007; McNeill, Lizotte, Krajcik, & Marx, 2006; Osborne, Erduran, & Simon, 2004).

Lastly, we wish to discuss some design implications for evidence choice in tasks that require students to develop or evaluate models. In the first task of the assessment, most students evaluated their models as being good because they addressed all the evidence, which was the case. There were no pieces of evidence that compelled students to provide mechanistic explanations at the molecular level. This was by design; we wanted to see if students would be able to come up with such mechanisms on their own and based on what they had learned in the unit. This expectation may have been too ambitious and we may have seen better models had one of the evidence pieces mentioned a protein or a genetic mutation, thereby cuing students to think about the molecular entities they had studied. The implication we draw here is that it may be necessary to provide evidence that relates to the explanatory mechanism one wants students to generate. This may seem fairly obvious, however, there is a tradeoff between providing evidence that "gives away" the answer, and providing evidence that directs students towards the appropriate grain size and nature of the desired explanation. A related tradeoff was also evident in the design of evidence 5 in the second part of the task. The evidence was somewhat ambiguous and open to multiple interpretations and students capitalized on this property and used the evidence in support of the incorrect model. Thus, there is also a delicate balance between choosing evidence that is too clear cut and that hides the under-determined nature of most real-world evidence-model relationships, and providing evidence that invites alternative interpretation and does not clearly rule out the erroneous model. We do not have guidelines or solutions to address these tradeoffs, but we believe the field would benefit from more explicit discussion of the design challenges intrinsic to engaging students with authentic disciplinary practices.

References

Chinn, C.A., and L.A. Buckland. 2012. Model-based instruction: Fostering change in evolutionary conceptions and in epistemic practices. In Evolution challenges: Integrating research and practice in teaching and learning about evolution, ed. K.S. Rosengren et al., 211–232. New York: Oxford University Press.

Dianovky, M., Duncan, R.G., & Chinn, C. A. (2013). Using Evidence to evaluate multiple competing models. Paper presented at the annual meeting of the *National Association of Research in Science Teaching.* Rio Grande, Puerto Rico.

Doerr, H., & Lesh, R. (2003). Foundations of a model and modeling perspective on mathematics teaching, learning, and problem solving. *In Beyond Constructivism*, pp. 3-33.

Driver, R., & Leach, J. M. R., Scott, P.(1996). *Young people's images of science*. Open University Press, Buckingham

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. Journal of Research in Science Teaching, 44(7), 938-959.
Duncan, R. G., & Tseng, K. A. (2011). Designing project-based instruction to foster generative and mechanistic understandings in genetics. *Science Education,* 95(1), 21-56.
Duncan, R. G., Freidenreich H. B., Chinn C. A., & Bausch, A. (2011). Promoting middle school students' understanding of molecular genetics. *Research in Science Education,* 41(2), 147-167.
Duschl, R. (2007). Quality argumentation and epistemic criteria. In S. Erduran, & M. Jiménez-Aleixandre (eds.), *Argumentation in science education: perspectives from classroom-based research* (pp. 159-175). Amsterdam: Springer.
Giere, R.N. (2004). How models are used to represent reality. *Philosophy of Science*, 71, 742–752.
Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21, 725–740.
Grosslight, L., Unger, C., Jay, E., & Smith, C.L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28, 799–822.
Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.
Kuhn, T.S. (1977). *The essential tension*. Chicago: University of Chicago Press.
Latour, B. (1987). *Science in Action*, Cambridge, MA: Harvard University Press.
Lehrer, R., & Schauble. L. (2000). Modeling in mathematics and science. In R. Glaser (Ed.), *Advances in instructional psychology: Vol 5*. *Educational design and cognitive science* (pp. 101-159). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Lewis, J., & Kattmann, U. (2004). Traits, genes, particles and information: re-visiting students' understandings of genetics. *International Journal of Science Education*, 26(2), 195-206.
Lewis, J., &Wood-Robinson, C. (2000). Genes, chromosomes, cell division, and inheritance- do students see any relationship? *International Journal of Science Education*, 22, 177–195.
Longino, H. E. (2002). *The fate of knowledge*. Princeton, NJ: Princeton University Press.
Marbach-Ad, G. (2001). Attempting to break the code in student comprehension of genetic concepts. *Journal of Biological Education*, 35(4), 183–189.
Marbach-Ad, G., & Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, 34(4), 200–210.
McNeill, K.L., Lizotte, D.J., Krajcik, J., & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences*, 15, 153–191.
National Research Council (1996). *National science education standards*. Washington, DC: National Academies Press.
National Research Council (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.
NGSS Lead States. (2013). *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.
Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41(10)*, 994–1020.
Pluta, W. J., Chinn, C. A., Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching*, *48*, 486-511.
Schwarz, C., & White, B. (2005). Meta-modeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165-205.
Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., … & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners.*Journal of Research in Science Teaching*, *46*(6), 632-654.
Treagust, D. F., Chittleborough, G. D., & Mamiala, L. T. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education,* 24, 357-368.
van Mil, M. H., Boerwinkel, D. J., & Waarlo, A. J. (2013). Modelling molecular mechanisms: A framework of scientific reasoning to construct molecular-level explanations for cellular behaviour. Science & Education, 22(1), 93-118.
Venville, G. J., & Treagust, D. F. (1998). Exploring conceptual change in genetics using a multidimensional interpretive framework. Journal of Research in Science Teaching, 35, 1031–1055.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1008634. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Author Index

Pages 1-622: Volume 1
Pages 623-1176: Volume 2
Pages 1177-1764: Volume 3

Abbott, Robert D., 962, 1152
Abrahamson, Dor, 23, 1593
Acosta, Alisa, 673
Adams, Deanne, 1199
Aguilar, Stephen, 1665
Ahn, June, 174, 455, 657, 1719
Alcalá, Lucía, 13
Aleven, Vincent, 977, 1352
Alfonso-Gurneau, Jasmine, 12
Alibali, Martha W., 479, 649, 1042
Allert, Heidrun, 238
Alonzo, Alicia, 1037
Alston, Alice, 410
Alvarenga, Claire, 1012
Anderson, Emma, 118, 1456
Anderson, Janice, 1641
Andrade-Lotero, Alejandro, 1637
Arastoopour, Golnaz, 150, 1680
Arias, Anna Maria, 1426, 1749
Arici, Anna, 697
Asterhan, Christa S. C., 1342, 1684
Azevedo, Roger, 309, 1052
Bachfischer, Agnieszka, 1283
Baker, Ryan S., 222
Ballweber, Christy, 1647
Banerjee, Amartya, 1603
Bang, Megan, 4, 12, 1372, 1436
Bannister, Nicole, 1209
Barab, Sasha, 697
Barany, Amanda, 1199
Barber, Jacqueline, 1117
Barker, Lisa M., 1446
Barrientos, Kristina, 1012
Barron, Brigid, 1264
Barth, Armin, 1179
Barth-Cohen, Lauren, 325, 1531
Barzilai, Sarit, 721
Basu, Satabdi, 1097
Baumeister, Antonia E. E., 38
Beauvineau, Yves, 1022
Beck, Luisa, 1623
Bell, Alexander, 1082
Bell, Philip, 1228, 1426, 1710
Bemis, Carrie Allen, 1489
Ben-David Kolikant, Yifat, 1362
Ben-Zvi, Dani, 394, 1549, 1677
Bernstein, Debra, 1485
Berson, Eric, 1537
Bevan, Bronwyn, 1711
Bhatnagar, Sameer, 982
Bielaczyc, Katerine, 1315, 1677
Biemans, Harm J. A., 1569
Bientzle, Martina, 102
Birmingham, Daniel, 952
Biswas, Gautam, 1097, 1352
Black, John B., 230
Blair, Kristen P., 1179
Blikstein, Paulo, 863, 1147, 1669
Bolling, Amy, 1436
Bolzer, Markus, 1416
Boncoddo, Rebecca, 479, 649
Bonsignore, Elizabeth, 174, 455, 657
Booker, Angela N., 919
Borge, Marcela, 753
Boston, Melissa D., 997
Bouchet, François, 309, 1052
Bowker, Geoffrey C., 6
Boxerman, Jonathan, 1583
Brady, Corey, 1199, 1388, 1603
Bransford, John, 1647
Brennan, Karen, 18, 1559
Breuleux, Alain, 14
Briseño, Adriana, 879
Britt, M. Anne, 1541
Brodie, Karin,
Brooks, Christopher, 1691
Brown, Willard, 1571
Bryant, Julie, 1643
Buckingham Shum, Simon, 150, 1680
Burke, Jeff, 1436
Burke, Quinn, 86, 1219
Burkett, Candice, 1541
Burleson, Winslow, 278, 847
Buxton, Cory, 1332
Caccamise, Donna, 1002
Cadeiras, Martin, 1012
Caires, Roxane, 495
Cakir, Murat Perit, 1112
Calabrese Barton, Angela, 952
Caleon, Imelda, 535
Callanan, Maureen, 1228
Cantarero, Andrea, 1563
Capps, Daniel, 325, 1531
Carlone, Heidi, 1332
Carney, Michael, 1456
Cartier, Jennifer, 599, 1621
Cartun, Ashley, 348
Castillo, Tim, 1563
Cerratto Pargman, Teresa, 1597
Cervantes, Francisco, 1559
Chae, Hui Soo, 1709
Chaffee, Rachel, 1557

 © ISLS

Champney, Danielle, 62
Chan, Carol K. K., 126, 333
Chao, Jie, 1523
Charles, Elizabeth S., 982
Chase, Kiera, 23
Chen, Gaowei, 583
Chen, Vivian, 1669
Chen, Xiaodi, 1623
Chen, Ying-Chih, 641
Cheng, Britte Haugan, 1661
Cheng, Harry, 1609
Cheng, Julius, 1645
Chi, Michelene T. H., 847, 972, 1527
Chi, Min, 1645
Chin, Doris B., 1179
Ching, Cynthia Carter, 1273
Chinn, Clark A., 615, 1122, 1189, 1686
Chiu, Jennifer L., 1523
Cho, Young Hoan, 535
Choi, Gi Woong, 1067
Choi, Jinnie, 607
Choi, Sung-Youn, 1635
Chowning, Jeanne, 1426
Chu, Haiwen, 1717
Chung, Huy Q., 418
Clariana, Roy B., 1543
Clark, Douglas B., 1199, 1342, 1388, 1657
Clarke, Sherice N., 583, 1684
Clarke-Midura, Jody, 1731
Clase, Kari L., 1567
Clegg, Tamara, 174, 455
Clement, John J., 503
Clodfelter, Erika, 1052
Close, Eleanor W., 1533
Close, Hunter G., 1533
Cobb, Paul, 4
Cober, Rebecca, 14, 1273
Coffey, Janet E., 1406
Cole, Michael, 1254
Collier, Wesley, 1680
Collins, Allan, 1315
Collins, Jamie, 1625
Conforti Preszler, Noelle, 1647
Conlin, Luke D., 31
Conn, Jessica, 1533
Conner, Laura, 1555
Constantin, Ana-Maria, 1511
Cook, Melissa Sunshine, 625
Cooper, Stephen, 992
Coppens, Andrew, 13
Cordy, Michelle, 1521
Core, Mark, 1057
Corona Caraveo, Yolanda, 13
Correa-Chávez, Maricela, 13
Cox, Christopher, 1509
Cress, Ulrike, 102
Crooks, Noelle, 1042
Cruz, Daniel, 1012
Cunningham, Christine M., 1587
Cunningham, Jahneille, 1671
Curnow, Joe, 134, 206
Cutler, Christopher T., 839
D'Amico, Laura, 1362
D'Angelo, Cynthia, 1489, 1732
Dalvi, Tejaswini, 1565
Daly, Alan J., 426
Damşa, Crina, 440, 1283, 1733
Danielak, Brian A., 1047
Danish, Joshua A., 1273, 1323, 1637
Dasgupta, Chandan, 1497
Davenport, Jodi L., 1583
Davis, Elizabeth A., 1426
Dayton, Andy, 13
DeBarger, Angela Haydel, 1022, 1703
DeJaegher, Crystal J., 1523
DeLiema, David, 1750
DeSutter, Dane, 987, 1599
Dede, Chris, 1579, 1581
Deitrick, Elise, 591
Delen, Ibrahim, 947
Delgado, Cesar, 262
Dempsey, Mary, 12
Deng, Mario C., 1012
Denner, Jill, 1007
Derry, Sharon J., 370, 1315
DiGiacomo, Daniela K., 70, 729
DiSalvo, Betsy, 793
DiSalvo, Carl, 793
Dianovsky, Michael T., 816
Dillenbourg, Pierre, 15, 1017
Dingyloudi, Filitsa, 761, 1132
Dixon, Colin, 1591
Doane, William E. J., 1047
Dookie, Lesley, 402, 1751
Dow, Steven, 1515
Drake, Joel, 1617, 1659
Duarte Olson, Izabel, 1372
Duck, Jennifer, 947
Duckles, Joyce, 737
Dugdale, Michael, 982
Dukeman, Anton, 1097
Duncan Valentine, Keri, 745
Durik, Amanda M., 713
Duschl, Richard,
Dussault, Mary, 1511
Dutilly, Erik, 1302, 1527
Easterday, Matthew W., 317
Eberbach, Catherine, 1228
Eberle, Julia, 463, 1734
Edelson, Daniel C., 1466
Edwards, Ann R., 1406
Eilam, Billie, 937
Eisenberg, Michael, 190
Eitel, Karla Bradley, 1509
El Taraboulsi, Sherine, 1302
Elby, Andrew, 286, 1037, 1406
Elinich, Karen, 1219, 1456
Engelmann, Katharina, 246
Engelmann, Tanja, 1543, 1545
Engeström, Yrjö, 1254, 1308

Enyedy, Noel, 625, 1273, 1293, 1323, 1436
Escudé, Meg, 1474
Eshet-Alaklai, Yoram, 721
Esmonde, Indigo, 134
Etkina, Eugenia, 1501
Evans, Abigail, 1157
Evans, Sarah, 1633
Eylon, Bat-Sheva, 14
Faber, Lori, 12
Farris, Amy Voss, 1102
Fatolapour, Mitra, 1515
Feng, Mingyu, 527, 1022
Feyzi Begnagh, Reza, 309
Fields, Deborah A., 855, 927, 1219
Finnigan, Kara S., 426
Fischer, Frank, 246, 463, 1189, 1416, 1704
Fischer, Gerhard, 15, 190
Fitz-Gibbon, Taylor, 1436
Flood, Virginia J., 1593
Fogo, Bradley, 1362, 1446
Fong, Cresencia, 1273
Forbes, Cory T., 46, 942
Ford, Michael J., 1698
Forman, Ellice A.,
Forsyth, Carol, 1611
Franz, Paul, 1137
Fricchione, Cheryl, 1631
Friend, Angela, 1002
Fritts, Jeremy, 1661
Fu, Ella, 333
Fuhrmann, Tamar, 863
Fujii, Keita, 967
Funk, Alexandra, 1416
Galbreath, Marcie, 1653
Gallagher, Daniel J., 12
Gallegos Greenwich, Joel, 737
Garcia, Iolanda, 1072
Gaus, Katharina, 1179
Gaydos, Matthew, 1199
Geller, Benjamin D., 1142
Gentry, Christine, 1587
Georgiou, Yiannis, 1577, 1752
Gerard, Libby, 386
Gerber, Elizabeth M., 317
Gibbons, Thor, 1342
Gilbuena, Debra, 1651
Ginsberg, Daniel, 1723
Girotto, Victor, 278
Glogger, Inga, 1179
Gnoli, Alessandro, 1273
Gobert, Janice, 222
Goggins, Sean, 753, 1535
Golan Duncan, Ravit, 158, 607, 615, 1122, 1189
Golbeck, Susan, 1491
Goldberg, Benjamin, 1352
Goldberg, Tsafrir, 1362
Goldin, Simona, 1505
Goldman, Shelley, 1228
Goldman, Susan R., 1541, 1571
Goldwasser, Lloyd, 689
Gomez, Kimberley, 1671, 1700
Gomez Zaccarelli, Florencia, 1724
Gong, Bowyee, 1661
Gonzalez, Carlos, 1315
Gonzalez, Martha Rocío, 1247
González-Howard, Maria, 1499
Goodell, Alexandra, 962
Goodman, Alyssa A., 1511
Goodyear, Peter, 1283
Gordon, Andrew, 1057
Gottlieb, Eli, 1247
Gotwals, Amelia Wenk, 1503
Gouvea, Julia S., 705, 1142
Graesser, Arthur C., 1611
Graville-Smith, Cynthia, 1436
Greenhow, Christine, 1342
Greenleaf, Cynthia, 1571
Greeno, James G., 1396
Gresalfi, Melissa, 1293, 1396
Gressick, Julia, 370, 1735
Griffin-EL, Nosakhere, 1302
Groneman, Christine, 1002
Grotzer, Tina, 1579, 1581
Grover, Shuchi, 992
Gubbels, Michael, 174, 455
Guerra, Paulo, 1273
Guidalli, Barbara, 1302
Guillemette, Jonathan, 982
Guloy, Sheryl, 1362
Gupta, Ayush, 286, 1406, 1605, 1607
Guthrie, Mareca, 1555
Gutierrez, Pedro J., 972
Gutiérrez, Kris D., 4, 13, 70, 729, 1254, 1293, 1382, 1711
Hackbarth, Alan J., 1315
Hall, Rogers, 3, 1237, 1302, 1308, 1323, 1396
Hallinen, Nicole R., 31, 1179, 1645
Halverson, Erica Rosenfeld, 839, 1436
Halverson, Kristy L., 1567
Hammer, David, 633
Hand, Brian, 641
Hand, Victoria, 1209, 1709
Hanny, Courtney, 737
Harding, Tricia, 1711
Harris, Christopher, 1489, 1703
Harrison, Simon, 1655
Hayes, Elisabeth, 1264
Headrick Taylor, Katie, 1237, 1264, 1736
Heck, Tamara, 1022
Heffernan, Neil T., 527
Helding, Brandon, 1527
Heredia, Sara, 1727
Herrenkohl, Leslie Rupert, 1152, 1247
Hertel, Jonathan D., 1587
Hetmanek, Andreas, 1189
Hever, Rakheli, 1342
Heyd-Metzuyanim, Einat, 1396, 1639
Hickey, Daniel T., 11, 1027
Hirsh, Alon, 1372
Hjorth, Arthur, 1553, 1603

Hmelo-Silver, Cindy E., 11, 410
Hoadley, Christopher, 1466, 1698
Hod, Yotam, 1519, 1549, 1677
Hokayem, Hayat, 1503
Holbert, Nathan, 1199, 1388
Holmqvist Olander, Mona, 1087
Honey, Rose, 1466, 1653
Honwad, Sameer, 1466, 1653, 1718
Hooper, Paula K., 919
Horn, Ilana Seidel, 1209
Horn, Michael, 1603
Horstman, Theresa, 1647
Horwitz, Paul, 681
Hougham, R. Justin, 1509
Howley, Iris, 1561
Huang, Jason, 1609
Hug, Sarah, 214
Hung, Leah C.-C., 1122
Hwang, Soyeon, 1515
Iiskala, Tuike, 1651
Ilomäki, Liisa, 1283
Irving, Paul W., 1513
Itow, Rebecca, 14
Jackson, G. Tanner, 1481
Jackson, Jonathan, 1511
Jackson, Kara, 4
Jacobson, Michael J., 362
Jain, Rishika, 23
James, Katherine, 1571
Jamshidi, Arash, 705
Jensen, Bryant, 903
Jermann, Patrick, 1017
Jiang, Yang, 222
Jimenez, Osvaldo, 665
Jimenez Pazmino, Priscilla, 198, 1273
Jin, Hui, 823
Jipson, Jennifer, 1228
Jocuns, Andrew, 1692
Johnson, Angela, 1332
Johnson, Raymond, 1127, 1171, 1254
Johnson-Glenberg, Mina C., 1199
Jones, Kimberly, 737
Jordan, Michelle E., 1166
Joyce-Gibbons, Andrew, 1737
Judele, Raluca, 1342
Jurow, A. Susan, 1254, 1302
Kafai, Yasmin B., 86, 855, 1219, 1388
Kahn, Jennifer, 1649
Kalaitzidis, TJ, 839
Kali, Yael, 14
Kamarainen, Amy, 1579, 1581
Kane, Britnie Delinger, 1209
Kang, Raymond, 911
Kang, Sin-Hwa, 1057
Kapon, Shulamit, 887, 1062
Kapur, Manu, 362
Ke, Li, 182
Keifert, Danielle, 1753
Kelton, Molly, 1396
Kern, Anne, 1653
Kessler, Aaron M., 599, 997, 1621
Khan, Saadia A., 230
Khattak, Laaraib, 1521
Killingsworth, Stephen, 1199, 1388
Kim, Nam Ju, 1617
Kim, Stephanie, 1573
Kimmerle, Joachim, 102
King, Whitney L., 927
Kinnebrew, John S., 1097, 1352
Kintsch, Eileen, 1002
Kirschner, Paul A., 1569, 1680
Kirshner, Ben, 348, 1302, 1711
Klahr, David, 1189
Klein, Elaine, 1426
Klein, Valerie, 1631
Klopfer, Eric, 118, 1388
Knight, Daniel W., 1601
Knight, Simon, 150, 1680
Ko, Mon-Lin Monica, 54, 1571
Koedinger, Kenneth R., 1352
Koehler, Jessica, 118
Koh, Elizabeth, 535
Kolodner, Janet,
Kolodziej, Richard, 1543
Kong, Fan, 1474
Kooser, Ara, 1625
Kopcha, Theodore J., 745
Koretsky, Milo, 1651
Koschmann, Timothy, 551, 1323
Kosonen, Kari, 1283
Kotsopoulos, Donna, 1521
Kotys-Schwartz, Daria A., 1539, 1601
Kozlov, Michail D., 1543, 1545
Krajcik, Joseph, 1635, 1703
Krehbiel, Matt D., 1711
Krinks, Kara, 1388
Krishnan, Gokul, 1547
Krist, Christina, 270
Krug O'Neill, Michaela, 1505
Kuhn, Alex, 947
Kumar, Melissa, 1293
Kumar, Vishesh, 23
Kumpulainen, Kristiina, 18
Kuo, Annie Camey, 1663
Kuo, Eric, 62
Kutsson, Ola, 1597
Kyza, Eleni A., 11, 1577
König, Katrin, 1545
Könings, Karen, 14
Körndle, Hermann, 1416
Lachapelle, Cathy P., 1587
Laferrière, Therese, 14
Lai, Kevin, 769
Lajoie, Susanne, 1352
Lakkala, Minna, 1283
Lam, Diane P., 689
Lam, Rachel J., 1667
Land, Susan M., 378, 1067
Langemeyer, Melanie, 1521
Langer-Osuna, Jennifer, 1293

Lanouette, Kathryn, 1537
Lapshin, Yuri, 1308
Larson, Joanne, 737
Lasry, Nathaniel, 982
Lauff, Carlye, 1539
Laumatia, Laura, 1653
Lauricella, Alexis R., 1264
Lave, Jean, 3
Law, Nancy, 777
Leary, Heather, 110, 1127, 1171, 1254
Lee, Carol D., 1382
Lee, Hee-Sun, 681
Lee, June H., 1264
Lee, May, 182
Lee, Nathanie, 962
Lee, Tiffany R.,
Lee, Ung-Sang, 1671
Lee, Victor R., 110, 1032, 1273, 1617, 1659
Lee, Yeung, 777
Lehrer, Richard, 1315
Leiba, Moshe, 1416
Lembke, Julia, 238
Lenton, Kevin, 982
Leones, Tiffany, 1661
Levin, Mariana, 1323
Levinson, Amber, 1228, 1264
Levy, Sharona T., 1372
Lewis, Armanda, 1466
Lewis, Colleen, 495
Lewittes, Becky, 174, 455
Liberti, Jamie, 1491
Light, Gregory J., 1573
Lin, Emily S., 1725
Lin, Feng, 126
Lindgren, Robb, 341, 1199, 1456
Linn, Marcia C., 14, 386, 471, 769, 895, 1585
Lira, Matthew, 987, 1595
Little, Angela, 1551
Littleton, Karen, 150
Littrell-Baez, Megan, 1002
Litts, Breanne K., 1726
Litvack, Elyse, 1228
Livingston, Ayesha, 1625
Lo, Abraham S., 1754
Lo, Stanley M., 1573
Lo, Wan-Tzu, 947
Lockwood, Elise, 1042
Loibl, Katharina, 1179
Long, Yanjin, 1352
Lonn, Steven, 1665, 1691
Loper, Suzanna, 1117
Lotas, Sasha, 1162
Lotridge, Erin, 1511
Lozano, Cecil, 278
Luce, Megan, 1228
Lucero, Margaret, 262
Ludvigsen, Sten, 1283
Luehmann, April, 1426, 1711
Lui, Michelle, 301, 673
Luna, Melissa J., 1406
Lyons, Cheryl A., 354
Lyons, Leilah, 198, 1273, 1456, 1709
Lyu, Siwei, 823
López, Angélica, 13
López Silva, Brenda A., 198, 1273
Ma, Guanzhong, 78
Ma, Jasmine Y., 1237, 1396
Macalalag Jr., Augusto Z., 559
Madsen, Adrian, 1082
Mafa, Moseli A., 1529
Magana, Alejandra J., 1738
Magnifico, Alecia Marie, 839, 1219
Maher, Carolyn A., 410
Makela, Marja-Liisa, 1651
Mancevice, Nicole, 1671
Mandl, Heinz, 1483
Mangen Sypher, Ofelia, 1466
Manz, Eve, 575, 1739
Marchand, Gwen C., 1487
Marin, Ananda, 12, 1237, 1372
Markauskaite, Lina, 14, 1283
Marlino, Mary,
Martin, Andrew, 1669
Martin, Caitlin Kennedy, 1264
Martin, Lee, 1591
Martin, Wendy, 1559
Martinez, Jacob, 1007
Martinez-Garza, Mario M., 1199, 1657
Martinuk, Mathew "Sandy", 1082
Mathis, Lele, 1605
Matuk, Camillia, 14, 895
Mazziotti, Claudia, 1179, 1755
McCallie, Ellen, 1698
McClain, Lucy Richardson, 1067, 1228
McCormick, Mary, 633
McDonald, Scott, 1615
McEldoon, Katherine, 1179
McElhaney, Kevin, 386
McGee, Steven, 713, 947
McGinty, Megan, 1372
McGowan, Veronica, 1426
McKagan, Sarah B., 1082
McKay, Tabitha, 1563
McKenney, Susan, 14, 1700
McKinney de Royston, Maxine, 1382
McNamara, Danielle S., 15, 1416, 1481
McNeill, Katherine L., 1117, 1499
McWilliams, Jenna, 1692
Medin, Douglas L., 1372
Mejía-Arauz, Rebeca, 13
Mendelson, Adam, 689
Mendoza, Elizabeth, 1474
Menzer, Melissa, 1342
Metcalf, Shari, 1579, 1581
Metz, Kathleen E., 1537
Meyer, Chris, 1653
Meyer, Eric, 737
Meyer, Xenia S., 689
Michaelis, Joseph E., 1525
Mikkola, Anna, 18

Miller, Alison R., 354
Miller, Brant G., 1509
Miller, Susan, 1127, 1254
Milrad, Marcelo, 1597
Miyake, Naomi,
Moher, Tom, 198, 1273, 1497
Mohney, Michael R., 1067
Moore, Emily B., 1199
Moorthy, Savitha, 1489
Mor-Hagani, Shiri, 394
Morozov, Andrew, 1152
Moses, George, 737
Moses, Robert, 737
Mote, Chad, 1219
Mouza, Chrystalla, 1709
Mu, Jin, 333
Mudrick, Nicholas, 309, 1052
Mulder, Martin, 1569
Muldner, Kasia, 278, 847
Munter, Charles, 1396
Murphy, Robert, 527
Muukkonen, Hanni, 1283
Mørck, Line Lerche, 487
Nachlieli, Talli, 1639
Naik, Shweta, 1505
Namdar, Bahadir, 254
Narciss, Susanne, 1416
Nasir, Na'ilah Suad, 1382, 1436
Nathan, Mitchell J., 479, 649, 1042, 1525
Natriello, Gary, 1709
Nerland, Monika, 1283
Neumeister, Michael, 551
Niedermeier, Sandra, 1483
Nielsen, Rodney D., 1527
Nissen, Jayson, 142
Nolen, Susan Bobbitt, 962, 1651
Noroozi, Omid, 1569
Nourian, Saeid, 1517
Novik, Tamar, 1549
Nsair, Ali, 1012
Nussbaum, E. Michael, 1487
O'Connell, Brian, 591
O'Connor, Kevin, 6, 1539, 1692
O'Neill, D. Kevin, 1254, 1362
Ogan, Amy, 1740
Olander, Clas, 1087
Olson, Arthur J., 1583
Osborne, Jonathan, 1189
Oshima, Jun, 967
Oshima, Ritsuko, 967
Otero, Nuno, 1597
Otero, Valerie, 567, 800
Packer, Martin J., 1247, 1254
Palincsar, Annemarie S., 1426
Palius, Marjory F., 410
Pallant, Amy, 681
Panagiotopoulos, Dimitrios, 1756
Paquette, Luc, 222
Pardos, Zachary A., 1691
Parikh, Tapan, 1589
Parnafes, Orit, 887, 1323
Passmore, Cynthia, 705
Pastor, Dena Ann, 713
Paugh, Patricia, 1507
Pauw, Daniel, 455
Pavlik, Jr., Philip, 1611
Pea, Roy D., 992
Pearman II, F. Alvin, 1493
Peck, Raphaela, 1117
Peddycord III, Barry, 1691
Peek-Brown, Deborah, 1635
Pellicone, Anthony, 657
Penuel, William R., 1254, 1264, 1710
Peppler, Kylie, 18, 1219
Pereira Querol, Marco, 1308
Perkins, Katherine K., 1199
Perritano, Anthony, 1273
Peterman, Tana, 1426
Peters, Olaf, 1416
Phillips, Nathan, 1237
Phillips, Rachel S., 1647
Pier, Elizabeth, 479, 649
Piergallini, Mario, 1515
Pilitsis, Vicky, 158
Pinkard, Nichole, 12
Poitras, Eric, 1352
Pollack, Sarah, 1362
Polman, Joseph L., 1436, 1446
Pompea, Stephen, 1555
Price, Emily, 348
Price, Jeremy, 1117
Prudent, Mirlanda E., 1619
Pugh, Priya, 1372
Puhl, Thomas, 1342
Puntambekar, Sadhana, 1315
Puttick, Gillian, 1485
Quan, Gina, 1607
Quintana, Chris, 947, 1416, 1456
Rafferty, Anna N., 386
Rahm, Jrène, 1332
Raia, Federica, 1012
Ramberg, Robert, 1597
Ramos, Grecia, 1012
Rappa, Natasha Anne, 1757
Rau, Martina A., 977
Recker, Mimi, 110, 1495
Rees Lewis, Daniel, 317
Reeve, Richard, 14
Reimann, Peter, 362
Reiser, Brian J., 270, 1700
Reisman, Avishag (Abby), 1362, 1446
Renga, Ian Parker, 511
Renken, Maggie, 1686
Renkl, Alexander, 1179
Renner, Nan, 1720
Rentschler, Mark E., 1539
Resnick, Elana, 1501
Resnick, Lauren B., 583
Rhodes, Catherine, 1247
Rhodes, Emily, 174, 455

Richards, Jennifer, 286, 1406
Richter, Christoph, 238
Rindermann, Heiner, 38
Rinhart, Ronald W., 1189, 1758
Rivera, Valeria M., 1012
Rivet, Ann E., 354
Riviere, Dominique, 134
Roberts, Jessica, 1456
Robertson, Amy, 1406
Rogoff, Barbara, 13, 1382
Roll, Ido, 879, 1352, 1691
Ronen, Michaela, 1416
Roque, Ricarose, 1219, 1623
Roschelle, Jeremy, 7, 527
Roscoe, Rod D., 972, 1416
Rosebery, Ann, 1436
Rosenberg, Joshua, 182
Ross, Danielle, 599
Ross, Laurie, 957
Ross, Michael, 567
Rosé, Carolyn Penstein, 1515, 1561, 1639
Rubin, Andee, 1273
Ruhl, Elisa, 238
Rummel, Nikol, 15, 977, 1179, 1416
Rusk, Natalie, 1219, 1623
Russ, Rosemary S., 1388, 1406
Ruvalcaba, Omar, 13
Ryoo, Kihyun, 1585
Sadler, Philip, 1511
Safronova, Maria, 1308
Sagy, Ornit, 14, 1519
Salehi, Shima, 863
Salman, Fariha H., 378
Samei, Borhan, 1611
San Antonio-Tunis, Christopher, 1587
Sandoval, William A., 1107, 1189, 1315, 1686
Sannino, Annalisa, 1254, 1308
Sao Pedro, Michael, 222
Sato, Elissa, 471
Sato, Takumi, 952
Savio-Ramos, Caroline, 1199
Sawtelle, Vashti, 1142
Sayre, Eleanor C., 1082, 1513
Scalone, Giovanna, 1152
Schaefer, Sara, 1273
Schalk, Lennart, 1179
Schank, Patricia, 1022
Schauble, Leona, 7
Schellens, Tammy, 166
Schenke, Katerina, 1027
Scherrer, Jimmy, 1741
Schneider, Alyse, 1593
Schneider, Emily, 1691
Schneider, Leslie, 1575
Schumacher, Ralph, 1179
Schwartz, Daniel L., 31, 1179, 1645
Schwartz, Lisa Hope, 70
Schwarz, Baruch B., 294, 1342
Schwarz, Christina V., 46, 182, 942
Scipio, Déana, 871, 1077, 1663
Searle, Kristin Anne, 1759
Seely, Brian J., 1067
Sefton-Green, Julian, 18
Segedy, James R., 1352
Sengupta, Pratim, 1097, 1102, 1199, 1388, 1547
Sengupta-Irving, Tesha, 1293
Severance, Samuel, 1127, 1171, 1254, 1302
Sezen-Barrie, Asli, 1742
Sfard, Anna, 7, 1247, 1293
Shaffer, David Williamson, 150, 1680
Shah, Niral, 495
Shane-Sagiv, Chava, 1446
Shapiro, R. Benjamin, 591, 1388
Sharma, Kshitij, 1017
Shaw, Ashley, 1691
Shea, Molly, 1302, 1474
Shemwell, Jonathan, 142, 325, 1531
Shen, Ji, 254
Sherin, Miriam, 1406
Sherwood, Carrie-Anne, 1489
Shim, Julie, 1515
Shin, Namsoo, 1635
Shouse, Andrew, 1711
Shumar, Wesley, 1631
Shutt, Kari, 1152
Siemens, George, 15
Sigley, Robert, 410
Silberglitt, Matt, 1583
Siler, Stephanie A., 1189
Silva, Glauco, 1627
Silverman, Jason, 1631
Simon, Erica, 479
Siyahhan, Sinem, 1228, 1264
Slakmon, Benzi, 294
Slattery, Brian, 198, 1273
Slotta, James D., 14, 301, 673, 831, 1273
Smith, Carmen Petrick, 1647
Smith, Jeremy, 737
Smith, Scott, 1617
Snow, Erica L., 1416, 1481
So, Hyo-Jeong, 14
Soep, Elisabeth, 18
Soffer-Vital, Shira, 1743
Sommerhoff, Daniel, 1704
Songer, Nancy,
Sorum, Michael, 4
Sottilare, Robert, 1352
Spikol, Daniel, 1597
Squire, Kurt, 1199
Stahl, Gerry, 1697
Stampfer Wiese, Elaine, 1352
Stanford, Tina, 1489
Stegmann, Karsten, 463
Stein, Mary Kay, 808, 997
Steinbock, Daniel, 785
Stephens, A. Lynn, 503
Stevens, Reed, 7, 1264, 1323
Stevens, Shawn, 1635
Stieff, Mike, 987, 1599
Stiles, Abigail, 1643

Strawhacker, Amanda, 1485
Strijbos, Jan-Willem, 11, 761, 1132, 1416
Subramaniam, Mega, 657
Sun, Weiyi, 823
Sunbury, Susan, 1511
Sutherland, LeeAnn, 1635
Sutherland, Scot, 1609
Suárez, Enrique, 800
Svihla, Vanessa, 14, 1563, 1625, 1643
Swarat, Su, 1573
Sylvan, Elisabeth, 1485
Tabach, Michal, 1639
Tabak, Iris, 1456
Takeuchi, Lori, 1264
Talbot, Robert, 1527
Tally, William, 1559
Tamminga, Kenneth, 1466
Tan, Edna, 952
Tan, Esther, 14
Tang, Kok-Sing, 434
Tate, Carol, 615
Tate, Erika, 1022
Taub, Michelle, 309, 1052
Teasley, Stephanie D., 1665
Teeters, Leah, 1302
Tekkumru Kisa, Miray, 808
Telhan, Orkan, 1219
Teplovs, Chris, 1575
Thiry, Heather, 214, 1007
Thomas, Jakita O., 1092
Thompson, Jessica, 1615
Thummaphan, Phonraphee, 1152
Tierney, Gavin, 543, 962, 1077, 1761
Tinker, Robert, 681
Tissenbaum, Mike, 831, 1273
Tonso, Karen L., 1332
Toom, Auli, 1283
Torralba, José Antonio, 1302
Tran, Cathy, 1027, 1722
Trickett, Susan, 1527
Truitt, Tim S., 12
Tsai, Janet Y., 1601
Tscholl, Michael, 341, 1456
Tsovaltzi, Dimitra, 1342
Turpen, Chandra, 1142
Tyson, Kersti, 1643, 1744
Tzadok, Eynav, 721
Tzou, Carrie, 1228, 1555
Udomprasert, Patricia, 1511
Umphress, Jessica, 1228
Urrieta, Luis, 13
Uzunosmanoğlu, Selin Deniz, 1112
Valcke, Martin, 166
Van Dusen, Ben, 567
Van Eaton, Grant, 1199
Van Horne, Katie, 1613
Van Wart, Sarah, 1589
Vanderhoven, Ellen, 166
Vargas, Penelope, 1686
Varner, Laura K., 1416
Vasudevan, Veena, 855
Vauras, Marja, 1651
Vea, Tanner, 1228
Villani, Alberto, 1627
Virkkunen, Jaakko, 1308
Visintainer, Tammie, 1673
Vitale, Jonathan M., 769
Vogel, Freydis, 1704
Volet, Simone, 1651
Voogt, Joke, 14
Vossoughi, Shirin, 919, 1382, 1474
Vye, Nancy, 1152, 1647
Vänninen, Irene, 1308
Waala, Jessica, 649
Wadholm, Robert R., 1535
Wagh, Aditi, 1603
Wagmister, Fabian, 1436
Walden, Von, 1509
Walker, Breya, 1611
Walker, Erin A., 278
Walkington, Candace, 479, 649
Walkoe, Janet, 1406
Wall, Steven, 1641
Walsh, Elizabeth, 1709, 1745
Wang, Catherine, 1057
Wang, Joyce, 94, 118, 1456
Warren, Beth, 7, 1436
Wartella, Ellen, 1264
Waugh, Amanda, 657
Wecker, Christof, 519, 1189
Wei, Yu, 535
Weidler-Lewis, Joanna, 1539, 1629
Weinberger, Armin, 1342
Weintrop, David, 1388
Welsh, Ashley, 879
Wendell, Kristen B., 1507, 1565
West, Richard E., 1721
White, Tobin, 1609
Whittaker, Chris, 982
Wichmann, Astrid, 1416
Wienberg, Christopher, 1057
Wilensky, Uri, 1372, 1388, 1553, 1603
Wilkerson-Jerde, Michelle Hoda, 1746
Williams, Caroline, 479, 649
Williams, Michelle, 1022
Williams, Robert F., 1655
Wilson, Nicholas, 448
Wingert, Kerri, 1426
Wink, Donald J., 816
Winne, Phil, 1352
Wise, Alyssa Friend, 1680
Wiser, Marianne, 7
Wobbrock, Jacob O., 1157
Worker, Steven, 1762
Worsley, Marcelo, 1147
Wortham, Stanton, 1247
Wright, Christopher G., 1507
Wylie, Ruth, 972, 1527
Xiao, Sihan, 1107, 1763
Xie, Charles, 1517

Xing, Wanli, 1535
Yakes Jimenez, Elizabeth, 1563
Ye, Lei, 1495
Yee, Nikki, 879
Yeo, Amelia, 1042
Yip, Jason, 174, 455, 657
Yoon, Irene, 1209
Yoon, Susan, 94, 118, 1456
York, Adam, 348
Yuan, Min, 1495, 1617
Yuen, Johnny, 777
Zangori, Laura, 46, 942
Zarate, Rosalia Chavez, 1264
Zemel, Alan, 551
Zhang, Jianwei, 823
Zhang, Zhihui Helen, 1511, 1517
Zielezinski, Molly B., 1137
Zimmerman, Heather Toomey, 378, 1067, 1228
Zimmerman, Randi M., 1122
Zisk, Robert, 1501
de Saint-Georges, Ingrid, 1692
diSessa, Andrea A., 1323
ter Vrugte, Judith, 1760
van Aalst, Jan, 11, 78, 126, 333
van Es, Elizabeth A., 418, 1209, 1406
von Davier, Alina, 11